File copy

PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification [6] :		(11) International Publication Number: **WO 96/11951**
C07K 14/50, C12N 15/12, A61K 38/18	A2	(43) International Publication Date: 25 April 1996 (25.04.96)

(21) International Application Number: · PCT/US95/13075

(22) International Filing Date: 12 October 1995 (12.10.95)

(30) Priority Data:
08/323,337   13 October 1994 (13.10.94)   US
08/487,825   7 June 1995 (07.06.95)   US

(60) Parent Application or Grant
  (63) Related by Continuation
    US   08/323,337 (CIP)
    Filed on   13 October 1994 (13.10.94)

(71) Applicant (for all designated States except US): AMGEN INC. [US/US]; Amgen Center, 1840 Dehavilland Drive, Thousand Oaks, CA 91320-1789 (US).

(72) Inventors; and
(75) Inventors/Applicants (for US only): CHEN, Bao-Lu [–/US]; Suite 2105, 6400 Christie Avenue, Emeryville, CA 94608 (US). ARAKAWA, Tsutomu [–/US]; 3957 Corte Cancion, Thousand Oaks, CA 91320 (US).

(74) Agents: ODRE, Steven, M. et al.; Amgen Inc., Amgen Center, 1840 Dehavilland Drive, Thousand Oaks, CA 91320-1789 (US).

(81) Designated States: AL, AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

(54) Title: KERATINOCYTE GROWTH FACTOR ANALOGS

(57) Abstract

Novel analogs of proteins of KGF are provided comprising a charge-change by the deletion or substitution of one or more of amino acid residues 41-154 of Figure 2 (amino acids 72-185 of SEQ ID NO:2). These analogs are more stable than the corresponding parent molecule KGF.

KERATINOCYTE GROWTH FACTOR ANALOGS

Field of the Invention

5   The present invention relates to recombinant
DNA technology and protein engineering. Specifically,
recombinant DNA methodologies have been applied to
generate polypeptide analogs of keratinocyte growth
factor (KGF), a potent mitogen of non-fibroblast
10   epithelial cell growth, wherein the analogs have
improved stability as compared to that of the parent
KGF.

Background

15

         The complex process of tissue generation and
regeneration is mediated by a number of protein factors
sometimes referred to as soft tissue growth factors.
These molecules are generally released by one cell
20   type and act to influence proliferation of other cell
types. (Rubin *et al*. (1989), *Proc. Nat'l. Acad. Sci.
USA*, 86:802-806). Some soft tissue growth factors are
secreted by particular cell types and influence the
proliferation, differentiation and/or maturation of
25   responsive cells in the development of multicellular
organisms (Finch *et al*. (1989), *Science*, 245:752-755).
In addition to their roles in developing organisms, some
are significant in the continued health and maintenance
of more mature systems. For instance, in mammals there
30   are many systems where rapid cell turnover occurs. Such
systems include the skin and the gastrointestinal tract,
both of which are comprised of epithelial cells.
Included within this group of soft tissue growth factors
is a protein family of fibroblast growth factors (FGFs).
35   There are currently eight known FGF family
members which share a relatedness among primary

- 2 -

structures: basic fibroblast growth factor, bFGF
(Abraham *et al.* (1986), *EMBO J.*, 5:2523-2528); acidic
fibroblast growth factor, aFGF (Jaye *et al.* (1986),
Science, 233:541-545); int-2 gene product, int-2
5 (Dickson & Peters (1987), *Nature*, 326:833); hst/kFGF
(Delli-Bovi *et al.* (1987), *Cell*, 50:729-737 and
Yoshida *et al.* (1987), *Proc. Natl. Acad. Sci. USA*,
84:7305-7309); FGF-5 (Zhan *et al.* (1988), *Mol. Cell.
Biol.*, 8:3487-3495); FGF-6 (Marics *et al.* (1989),
10 *Oncogene*, 4:335-340); keratinocyte growth factor (Finch
et al. (1989), *Science*, 24:752-755) and hisactophilin
(Habazzettl *et al.* (1992), *Nature*, 359:855-858).

Among the FGF family of proteins, keratinocyte
growth factor ("KGF") is a unique effector of non-
15 fibroblast epithelial (particularly keratinocyte) cell
proliferation derived from mesenchymal tissues. The
term "native KGF" refers to a natural human (hKGF) or
recombinant (rKGF) polypeptide (with or without a signal
sequence) as depicted by the amino acid sequence
20 presented in SEQ ID NO:2 or an allelic variant thereof.
[Unless otherwise indicated, amino acid numbering for
molecules described herein shall correspond to that
presented for the mature form of the native molecule
(i.e., minus the signal sequence), as depicted by amino
25 acids 32 to 194 of SEQ ID NO:2.]

Native KGF may be isolated from natural human
sources (hKGF) or produced by recombinant DNA techniques
(rKGF) (Finch *et al.* (1989), *supra*; Rubin *et al.* (1989),
supra; Ron *et al.* (1993), *The Journal of Biological
30 Chemistry*, 268(4):2984-2988; and Yan *et al.* (1991), *In
Vitro Cell. Dev. Biol.*, 27A:437-438).

It is known that native KGF is relatively
unstable in the aqueous state and that it undergoes
chemical and physical degradation resulting in a loss of
35 biological activity during processing and storage (Chen
et al. (1994), *Pharmaceutical Research*, 11:1582-1589).

Native KGF is prone also to aggregation at elevated
temperatures and it becomes inactivated under acidic
conditions (Rubin et al. (1989), Proc. Natl. Acad. Sci.
USA, 86:802-806). Aggregation of native KGF in aqueous
5   solution also results in inactivated protein. This is
disadvantageous because such loss of activity makes it
impractical to store aqueous formulations of native KGF
proteins for extended periods of time or to administer
the protein over extended periods. Moreover, this is
10   particularly problematic when preparing pharmaceutical
formulations, because aggregated proteins have been
known to be immunogenic (Cleland et al. (1993), Crit.
Rev. Therapeutic Drug Carrier Systems, 10:307-377;
Robbins et al. (1987), Diabetes, 36:838-845; and
15   Pinckard et al. (1967), Clin. Exp. Immunol., 2:331-340).
Recombinant DNA technology has been utilized
to modify the sequences of various FGF family members.
For example, bFGF and aFGF have been modified by
deleting or substituting positively-charged residues,
20   which are important for heparin binding with neutral or
negatively-charged amino acids. It was reported that
the modified molecules resulted in reduced heparin
binding activity. Accordingly, it was taught that the
amount of modified molecule sequestered by heparin
25   and/or heparin-like molecules in a patient would be
reduced, thereby increasing potency as more of the FGF
will reach its targeted receptor (EP 0 298 723).
In order to improve or otherwise alter one or
more of the characteristics of native KGF, protein
30   engineering may be employed. Ron et al. (1993), J.
Biol. Chem., 268(4):2984-2988 reported modified KGF
polypeptides having 3, 8, 27, 38 or 49 amino acids
deleted from the N-terminus. Those polypeptides missing
3, 8, or 27 N-terminal residues retained heparin binding
35   ability; the others did not. Also, the polypeptides
missing 3 and 8 residues were reported as being fully

active, whereas the form missing 27 residues was 10-20 fold less mitogenic, and the forms lacking 38 or 49 amino acids did not have mitogenic activity. The stability of the modified KGF polypeptides was not
5   discussed or otherwise reported.

Published PCT application no. 90/08771, *supra*, also reported the production of a chimeric protein wherein about the first 40 N-terminal amino acids of mature form of native KGF were combined with the C-
10   terminal portion (about 140 amino acids) of aFGF. The chimera was reported to target keratinocytes like KGF, but it lacked susceptibility to heparin, a characteristic of aFGF but not KGF. The stability of the chimera was not discussed or otherwise reported.

15   Thus, the literature has not reported a modified KGF molecule having significantly improved stability relative to native KGF. Moreover, the literature has not reported sufficient teachings or evidence to provide a reasonable expectation of
20   successfully generating KGF molecules with such desirable characteristics.

It is not currently possible to predict the characteristics of a protein based upon the knowledge of only its primary structure. For example, the mitogenic
25   activity of aFGF is substantially increased in the presence of heparin, but the mitogenic activity of bFGF in the presence of heparin is only minimally increased, despite the fact that heparin tightly binds to bFGF [(Burgess and Maciag (1989), *Annu. Rev. Biochem.*,
30   58:575-606; Schreiber, *et al.* (1985), *Proc Natl. Acad. Sci. USA*, 82:6138-6142; and Gospodarowizc and Cheng (1986), *J. Cell Physiol.*, 128:475-485); and PCT 90/00418)]. In contrast, thymidine incorporation by BALB/MK cells is inhibited when heparin is included with
35   KGF in the culture medium.

- 5 -

Generally, the effects upon biological activity of any amino acid change upon the protein will vary depending upon a number of factors, including the three-dimensional structure of the protein and whether

5   or not the modification is to either the heparin binding region or the receptor binding region on the primary sequence of the protein. As neither the three-dimensional structure nor the heparin binding region and the receptor binding region on the primary sequence of

10   native KGF has been published, the knowledge within the art does not permit generalization about the effects of amino acid modifications to native KGF based upon the effects of amino acid modifications on even commonly categorized proteins.

15   It is the object of this invention to provide polypeptide analogs of KGF and nucleic acid molecules encoding such analogs that exhibit enhanced stability (e.g., when subjected to typical pH, thermal and/or other storage conditions) as compared to native KGF.

20

Summary of the Invention

The present invention provides novel, biologically active polypeptide analogs of KGF. For

25   purposes of this invention, the term "KGF" includes native KGF and proteins characterized by a peptide sequence substantially the same as the peptide sequence of native KGF which retain some or all of the biological activity of native KGF, particularly non-fibroblast

30   epithelial cell proliferation. By "characterized by a peptide sequence substantially the same as the peptide sequence of native KGF" is meant a peptide sequence which retains residues corresponding to Arg^{41}, Gln^{43}, Lys^{55}, Lys^{95}, Asn^{137}, Gln^{138}, Lys^{139}, Arg^{144}, Lys^{147},

35   Gln^{152}, Lys^{153} and Thr^{154} of SEQ ID NO:2 and which is encoded by a DNA sequence capable of hybridizing to

nucleotides 201 to 684 of SEQ. ID. NO:1, pr ferably
under stringent hybridization conditions.

The determination of a corresponding amino
acid position between two amino acid sequences may be
5   determined by aligning the two sequences to maximize
matches of residues including shifting the amino and/or
carboxyl terminus, introducing gaps as required and/or
deleting residues present as inserts in the candidate.
Database searches, sequence analysis and manipulations
10   may be performed using one of the well-known and
routinely used sequence homology/identity scanning
algorithm programs (e.g., Pearson and Lipman (1988),
Proc. Natl. Acad. Sci. U.S.A., 85:2444-2448; Altschul *et
al.* (1990), *J. Mol. Biol.*, 215:403-410; Lipman and
15   Pearson (1985), *Science*, 222:1435 or Devereux *et al.*
(1984), *Nuc. Acids Res.*, 12:387-395).

Stringent conditions, in the hybridization
context, will be stringent combined conditions of salt,
temperature, organic solvents and other parameters
20   typically controlled in hybridization reactions.
Exemplary stringent hybridization conditions are
hybridization in 4 X SSC at 62-67° C., followed by
washing in 0.1 X SSC at 62-67° C. for approximately an
hour. Alternatively, exemplary stringent hybridization
25   conditions are hybridization in 45-55% formamide, 4 X
SSC at 40-45°C. [*See*, T. Maniatis et. al., *Molecular
Cloning* (A Laboratory Manual); Cold Spring Harbor
Laboratory (1982), pages 387 to 389].

Thus, the proteins include allelic variations,
30   or deletion(s), substitution(s) or insertion(s) of amino
acids, including fragments, chimeric or hybrid molecules
of native KGF. One example of KGF includes proteins
having residues corresponding to Cys[1] and Cys[15] of SEQ ID
NO:2 replaced or deleted, with the resultant molecule
35   having improved stability as compared with the parent
molecule (as taught in commonly owned U.S.S.N.

08/487,825, filed on July 7, 1995). Specifically
disclosed molecules include: C(1,15)S, a KGF having
substitutions of serine for cysteine at amino acid
positions 1 and 15; ΔN15-ΔN24, KGFs having a deletion of
5   any one of from the first 15 to 24 amino acids of the N-
terminus of native KGF; ΔN3/C(15)S, a KGF having a
deletion of the first 3 amino acids of the N-terminus of
native KGF and a substitution of serine for cysteine at
amino acid position 15; ΔN3/C(15)-, a KGF having a
10   deletion of the first 3 amino acids of the N-terminus of
native KGF and a deletion of cysteine at amino acid
position 15; ΔN8/C(15)S, a KGF having a deletion of the
first 8 amino acids of the N-terminus of native KGF and
a substitution of serine for cysteine at amino acid
15   position 15; ΔN8/C(15)-, a KGF having a deletion of the
first 8 amino acids of the N-terminus of native KGF and
a deletion of cysteine at amino acid position 15;
C(1,15,40)S, a KGF having a substitution of serine for
cysteine at amino acid positions 1, 15 and 40;
20   C(1,15,102)S, a KGF having a substitution of serine for
cysteine at amino acid positions 1, 15 and 102; and
C(1,15,102,106)S, a KGF having a substitution of serine
for cysteine at amino acid positions 1, 15, 102 and 106.
      Another example of KGF includes proteins
25   generated by substituting at least one amino acid having
a higher loop-forming potential for at least one amino
acid within a loop-forming region of Asn^{115}-His^{116}-
Tyr^{117}-Asn^{118}-Thr^{119} of native KGF (as taught in
commonly owned U.S.S.N. 08/323,473, filed on October 13,
30   1994), specifically including H(116)G, a KGF having a
substitution of glycine for histidine at amino acid
position 116 of native KGF.
      A still further example includes proteins
having one or more amino acid substitutions, deletions
35   or additions within a region of 123-133 (amino acids

154-164 of SEQ ID NO:2) of native KGF; these proteins may have agonistic or antagonistic activity.

Surprisingly, it has been discovered that by deleting or substituting neutral or negatively charged
5 peptides for the more positively charged residues (i.e., substituting negatively charged residues for neutral or positively charged residues, or neutral residues for positively charged residues) of a KGF molecule (i.e., parent molecule), the resultant KGF analog has improved
10 stability as compared to the parent molecule. Preferably, in addition to having increased stability, the invention is directed to those analogs which also exhibit full biological activity (i.e., at least substantially similar receptor binding or affinity) as
15 compared to native KGF.

In another aspect of the invention, purified and isolated nucleic acid molecules encoding the various biologically active polypeptide analogs of KGF are described. In one embodiment, such nucleic acids
20 comprise DNA molecules cloned into biologically functional plasmid or viral vectors. In another embodiment, nucleic acid constructs may then be utilized to stably transform a procaryotic or eucaryotic host cell. In still another embodiment, the invention
25 involves a process wherein either a procaryotic (preferably *E. coli*) or eucaryotic host cell stably transformed with a nucleic acid molecule is grown under suitable nutrient conditions in a manner allowing the expression of the KGF analog. Following expression, the
30 resultant recombinant polypeptide can be isolated and purified.

A further aspect of the invention concerns pharmaceutical formulations comprising a therapeutically effective amount of a KGF analog and an
35 acceptabl pharmaceutical carrier. Such formulations

will be useful in treating patients afflicted with
epithelial diseases and injuries.

5   In this vein, another aspect relates to
methods of stimulating epithelial cell growth by
administering to a patient a therapeutically effective
amount of a KGF analog. In one embodiment, non-
fibroblast epithelial cells are the cells whose
proliferation is stimulated. Such epithelial cells
include various adnexal cells, pancreatic cells, liver
10   cells, and mucosal epithelium in the respiratory and
gastrointestinal tracts.

                    Brief Description of the Figures

15   Figure 1 shows the nucleotide (SEQ ID NO:1)
and amino acid (SEQ ID NO:2) sequences of native KGF
(the nucleotides encoding the mature form of native KGF
are depicted by bases 201 to 684 of SEQ ID NO:1 and the
mature form of KGF is depicted by amino acid residues 32
20   to 194 of SEQ ID NO:2).
           Figures 2A, 2B and 2C show the plasmid maps of
pCFM1156, pCFM1656 and pCFM3102, respectively.
           Figure 3 shows the nucleotide (SEQ ID NO:3)
and amino acid (SEQ ID NO:4) sequences of the construct
25   RSH-KGF.
           Figure 4 shows the nucleotide (SEQ ID NO:5)
and amino acid (SEQ ID NO:6) sequences of the construct
contained in plasmid KGF.
           Figure 5 shows the chemically synthesized
30   OLIGOs (OLIGO#6 through OLIGO#11; SEQ ID NO:12-17,
respectively) used to substitute the DNA sequence
between a KpnI site and an EcoRI site (from amino acid
positions 46 to 85 of SEQ ID No:6) in the construct
contained in plasmid KGF to produce the construct in
35   plasmid KGF(dsd).

Figure 6 shows the chemically synthesized OLIGOs (OLIGO#12 through OLIGO#24; SEQ ID NO:18-30, respectively) used to construct KGF (codon optimized).

5   Figure 7 shows the nucleotide (SEQ ID NO:31) and amino acid sequences (SEQ ID NO:32) of R(144)Q, a KGF analog having a substitution of glutamine for arginine at amino acid position 144 of native KGF.

Figure 8 shows the nucleotide (SEQ ID NO:33) and amino acid sequences (SEQ ID NO:34) of

10   C(1,15)S/R(144)E, a KGF analog having substitutions of serine for cysteine at amino acid positions 1 and 15 and a substitution of glutamic acid for arginine at amino acid position 144 of native KGF.

Figure 9 shows the nucleotide (SEQ ID NO:35)

15   and amino acid (SEQ ID NO:36) sequences of C(1,15)S/R(144)Q, a KGF analog having substitutions of serine for cysteine at amino acid positions 1 and 15 and a substitution of glutamine for arginine at amino acid position 144 of native KGF.

20   Figure 10 shows the nucleotide (SEQ ID NO:37) and amino acid (SEQ ID NO:38) sequences of ΔN23/R(144)Q, a KGF analog having a deletion of the first 23 amino acids of the N-terminus and a substitution of glutamine for arginine at amino acid position 144 of native KGF.

25   Figure 11 shows the amount of soluble protein, determined by size exclusion HPLC, as a function of incubation time at 37°C.

Figure 12 shows the estimated melting temperature (T_m) as a function of pH for native KGF,

30   C(1,15)S/R(144)Q and C(1,15)S/R(144)E.

Figure 13 shows a typical profile of mitogenic activity of R(144)Q, determined by measuring the incorporation of [^{3}H]-Thymidine during DNA synthesis and by comparing it to a native KGF standard curve.

35   Figure 14 shows a typical profile of the mitogenic activity of ΔN23/R(144)Q, determined by

measuring the incorporation of [^{3}H]-Thymidine during DNA synthesis and by comparing it to a native KGF standard curve.

Figure 15 shows a typical profile of the

5   mitogenic activity of C(1,15)S/R(144)Q, determined by measuring the incorporation of [^{3}H]-Thymidine during DNA synthesis and by comparing it to a native KGF standard curve.

Figure 16 shows a typical profile of the

10 mitogenic activity of C(1,15)S/R(144)E, determined by measuring the incorporation of [^{3}H]-Thymidine during DNA synthesis and by comparing it to a native KGF standard curve.

Figure 17 shows the nucleotide (SEQ ID NO:41)

15 and amino acid (SEQ ID NO:42) sequences of ΔN23/N(137)E, a KGF analog having a deletion of the first 23 amino acids of the N-terminus and a substitution of glutamic acid for asparagine at amino acid position 137 of native KGF.

20   Figure 18 shows the nucleotide (SEQ ID NO:43) and amino acid (SEQ ID NO:44) sequences of ΔN23/K(139)E, a KGF analog having a deletion of the first 23 amino acids of the N-terminus and a substitution of glutamic acid for lysine at amino acid position 139 of native

25 KGF.

Figure 19 shows the nucleotide (SEQ ID NO:45) and amino acid (SEQ ID NO:46) sequences of ΔN23/K(139)Q, a KGF analog having a deletion of the first 23 amino acids of the N-terminus and a substitution of glutamine

30 for lysine at amino acid position 139 of native KGF.

Figure 20 shows the nucleotide (SEQ ID NO:47) and amino acid (SEQ ID NO:48) sequences of ΔN23/R(144)A, a KGF analog having a del tion of the first 23 amino acids of the N-terminus and a substitution of alanine

35 for arginine at amino acid position 144 of native KGF.

Figure 21 shows the nucleotide (SEQ ID NO:49)

and amino acid (SEQ ID NO:50) sequences of ΔN23/R(144)L,
a KGF analog having a deletion of the first 23 amino
acids of the N-terminus and a substitution of leucine
for arginine at amino acid position 144 of native KGF.

5   Figure 22 shows the nucleotide (SEQ ID NO:51)
and amino acid (SEQ ID NO:52) sequences of ΔN23/K(147)E,
a KGF analog having a deletion of the first 23 amino
acids of the N-terminus and a substitution of glutamic
acid for lysine at amino acid position 147 of native
10   KGF.

Figure 23 shows the nucleotide (SEQ ID NO:53)
and amino acid (SEQ ID NO:54) sequences of ΔN23/K(147)Q,
a KGF analog having a deletion of the first 23 amino
acids of the N-terminus and a substitution of glutamine
15   for lysine at amino acid position 147 of native KGF.

Figure 24 shows the nucleotide (SEQ ID NO:55)
and amino acid (SEQ ID NO:56) sequences of ΔN23/K(153)E,
a KGF analog having a deletion of the first 23 amino
acids of the N-terminus and a substitution of glutamic
20   acid for lysine at amino acid position 153 of native
KGF.

Figure 25 shows the nucleotide (SEQ ID NO:57)
and amino acid (SEQ ID NO:58) sequences of ΔN23/K(153)Q,
a KGF analog having a deletion of the first 23 amino
25   acids of the N-terminus and a substitution of glutamine
for lysine at amino acid position 153 of native KGF.

Figure 26 shows the nucleotide (SEQ ID NO:59)
and amino acid (SEQ ID NO:60) sequences of
ΔN23/Q(152)E/K(153)E, a KGF analog having a deletion of
30   the first 23 amino acids of the N-terminus and a
substitution of glutamic acid for glutamine at amino
acid position 152 of native KGF and glutamic acid for
lysine at amino acid position 153 of native KGF.

Detailed Description

        In accordance with the present invention,
novel analogs of KGF are provided. The KGF analogs are
 5   produced by deleting or substituting one or more
specific, positively-charged residues in KGF.
        The KGF analogs have, among other properties,
an improved stability under at least one of a variety of
purification and/or storage conditions. For example,
10   the KGF analogs will generally be purified in a greater
yield of soluble, correctly folded protein. Moreover,
once the material is purified, it will be more stable to
pH, temperature, etc. as compared to the stability of
the parent molecule. As described in the Examples
15   section below (modified by the substitution of Gln and
Glu for arginine at position 144 [R(144)Q and R(144)E,
respectively] and in some instances modified at the N-
terminus as well) exhibit, relative to native KGF, (1) a
35 to 37.2 day increase of half-life upon storage at
20   37°C, (2) a 7.5-9.5% higher thermal melting temperatures
over the course of thermal unfolding, and (3) an
increase in T_m over a range of pH values.
        Although not intended to be bound by theory, a
possible reason for the enhanced stability of the
25   R(144)Q and R(144)E may be due to a reduction in overall
charge density of a cluster of basic residues, which is
inherently unstable due to charge repulsion, in the
absence of heparin. The results set forth below suggest
that the arginine residue at position 144 may correspond
30   to a residue in bFGF, as determined by X-ray
crystallography, which is reported to be within or near
a cluster of basic residues that mediate heparin binding
(Ago, et al. (1991), J. Biochem., 110:360-363; and
Erikss n et al. (1993), Protein Science, 2:1274-1284).
35   Native KGF contains 46 charged residues, 27 of
which carry a positive charge. In view of the results

obtained with the KGF analogs, a comparison of the
native KGF primary sequence with the primary sequence of
bFGF suggests that some of the 27 positively charged
residues form a cluster similar to a cluster found in
5 the tertiary structure of bFGF. Depending on the
location of such residues in the protein's three-
dimensional structure, substitution of one or more of
these clustered residues with amino acids carrying a
negative or neutral charge may alter the electrostatic
10 interactions of adjacent residues and may be useful to
achieve increased stability.

Thus other analogs, in addition to the
preferred R(144)Q specifically set forth herein, are
contemplated by the present invention. As used in this
15 invention, a "KGF analog" or a "polypeptide analog of
KGF" shall mean charge-change polypeptides wherein one
or more of amino acid residues 41-154 (amino acids 72-
185 of SEQ ID NO:2), specifically including amino acid
residues 123-133 (amino acids 154-164 of SEQ ID NO:2),
20 are deleted or substituted with a neutral residue or
negatively charged residue selected to effect a protein
with a reduced positive charge. Preferred residues for
modification are Arg^{41}, Gln^{43}, Lys^{55}, Lys^{95}, Lys^{128},
Asn^{137}, Gln^{138}, Lys^{139}, Arg^{144}, Lys^{147}, Gln^{152}, Lys^{153} or
25 Thr^{154}, with Gln^{138}, Lys^{139}, Arg^{144}, Lys^{147}, Gln^{152} or
Lys^{153} being more preferred and Arg^{144} being most
preferred. Preferred amino acids for substitution
include glutamic acid, aspartic acid, glutamine,
asparagine, glycine, alanine, valine, leucine,
30 isoleucine, serine and threonine, with glutamic acid,
glutamine, aspartic acid, asparagine and with alanine
being particularly preferred.

Any modification should give consideration to
minimizing charge repulsion in the tertiary structure of
35 the molecule; most preferably the analog will have
incr ased stability compared with the parent mol cule.

Obviously, the deletions or substitutions should not be
so numerous nor be made to residues of such close
proximity so as to set up charge repulsion between two
negatively-charged residues.

5   When the KGF analogs are biologically
generated, i.e., are the products of cellular expression
as opposed to the products of solid state synthesis,
proteolytic or enzymatic derivatization of naturally-
occurring products, etc., the nucleic acids encoding

10   such polypeptides will differ in one or more nucleotides
as compared to the native KGF nucleotide sequence. Such
nucleotides may be expressed and the resultant
polypeptide purified by any one of a number of
recombinant technology methods known to those skilled in

15   the art.
           DNA sequences coding for all or part of the
KGF analogs may include, among other things, the
incorporation of codons "preferred" for expression in
selected host cells (e.g., "E. coli expression codons");

20   the provision of sites for cleavage by restriction
enzymes; and the provision of additional initial,
terminal, or intermediate nucleotide sequences (e.g., as
an initial methionine amino acid residue for expression
in E. coli cells), to facilitate construction of readily

25   expressed vectors.
           The present invention also provides
recombinant molecules or vectors for use in the method
of expression of the polypeptides. Such vectors may be
comprised of DNA or RNA and can be circular, linear,

30   single-stranded or double-stranded in nature and can be
naturally-occurring or assemblages of a variety of
components, be they naturally-occurring or synthetic.
           Many examples of such expression vectors are
known. The components of the v ctors, e.g. replicons,

35   selection genes, enhancers, promoters, and th like, may
be obtained from natural sources or synthesized by known

- 16 -

procedures. In each case, expression vectors useful in
this invention will contain at least one expression
control element functionally associated with the
inserted nucleic acid molecule encoding the KGF

5   polypeptide analog. This control element is responsible
for regulating polypeptide expression from the nucleic
acid molecules of the invention. Useful control
elements include, for example, the *lac* system, the *trp*
system, the operators and promoters from phage λ, a

10   glycolytic yeast promoter, a promoter from the yeast
acid phosphatase gene, a yeast alpha-mating factor, and
promoters derived from adenovirus, Epstein-Barr virus,
polyoma, and simian virus, as well as those from various
retroviruses. However, numerous other vectors and

15   control elements suitable for procaryotic or eucaryotic
expression are known in the art and may be employed in
the practice of this invention.

Examples of suitable procaryotic cloning
vectors may include plasmids from *E. coli* (e.g. pBR322,

20   col El, pUC, and the F-factor), with preferred plasmids
being pCFM1156 (ATCC 69702), pCFM1656 (ATCC 69576) and
pCFM3102 (described in the Examples section, below).
Other appropriate expression vectors of which numerous
types are known in the art for mammalian, insect, yeast,

25   fungal and bacterial expression can also be used for
this purpose. The transfection of these vectors into
appropriate host cells can result in expression of the
KGF analog polypeptides.

Host microorganisms useful in this invention

30   may be either procaryotic or eucaryotic. Suitable
procaryotic hosts include various *E. coli* (e.g., FM5,
HB101, DH5α, DH10, and MC1061), *Pseudomonas*, *Bacillus*, and
Streptomyces strains, with *E. coli* being preferred.
Suitable eucaryotic host cells include yeast and other

35   fungi, insect cells, plant cells, and animal cells, such
as COS (e.g., COS-1 and COS-7) and CV-1 monkey cell

- 17 -

lines, 3T3 lines derived from Swiss, Balb-c or NIH
cells, HeLa and L-929 mouse cells, and CHO, BHK or HaK
hamster cells. Depending upon the host employed,
recombinant polypeptides produced in accordance herewith

5   will be glycosylated with mammalian or other eucaryotic
carbohydrates or may be non-glycosylated.

The preferred production method will vary
depending upon many factors and considerations; the
optimum production procedure for a given situation will

10 be apparent to those skilled in the art through minimal
experimentation. The resulting expression product may
then be purified to near homogeneity using procedures
known in the art. A typical purification procedure for
procaryotic cell production involves rupturing the cell

15 walls by high pressure or other means, centrifugation or
filtration to remove cellular debris, followed by ion
exchange chromatography of supernatant or filtrate and,
finally, hydrophobic interaction chromatography. If the
analog is expressed in insoluble form, another

20 purification technique involves first solublizing the
inclusion bodies containing the analogs followed by ion
exchange chromatography, then refolding of the protein,
and, finally, hydrophobic interaction chromatography.
Exemplary purification techniques are taught in commonly

25 owned U.S.S.N. 08/323,339, filed on October 13, 1994.
Generally, U.S.S.N. 08/323,339 teaches a method for
purifying a keratinocyte growth factor comprising: (a)
obtaining a solution comprising the KGF; (b) binding the
KGF from the solution of part (a) to a cation exchange

30 resin; (c) eluting the KGF in an eluate solution from
the cation exchange resin; (d) either passing the eluate
solution from part (c) through an appropriate molecular
weight exclusion matrix or performing hydrophobic
interaction chromatography on the eluate solution of

35 part (c); and (e) r covering the KGF from the molecular

weight exclusion matrix or hydrophobic interaction chromatography.

Of course, the analogs may be rapidly screened to assess their physical properties. The Examples sets

5   forth various well-known stability assays, although the specific assay used to test the analog is not critical. Moreover, the level of biological activity (e.g., receptor binding and/or affinity, mitogenic, cell proliferative and/or *in vivo* activity) may also be

10   tested using a variety of assays, some of which are set forth in the Examples section. Numerous assays are well-known and can be used to quickly screen the KGF analogs to determine whether or not they possess acceptable biological activity. One such assay

15   specifically tests the KGF analogs for the ability to bind to the KGF receptor (KGFR) by competing with ^{125}I-KGF binding (Bottaro *et al.* (1990), *J. Biol. Chem.*, 265:12767-12770; Ron *et al.* (1993), *J. Biol. Chem.*, 268:2984-2988). An alternative method for assaying

20   KGFR/KGF analog interactions involves the use of techniques such as real time biospecific interaction analysis (BIA) (Felder *et al.* (1993), *Molecular & Cellular Biology*, 13:1449-1455). Additionally a mitogenic assay can be utilized to test the ability of

25   the KGF analogs to stimulate DNA synthesis (Rubin *et al.* (1989), *supra*). Finally, cell proliferative assays can be utilized to test the ability of the KGF analogs to stimulate cell proliferation (Falco, *et al.* (1988), *Oncogene*, 2:573-578). Using any of the aforementioned

30   assay systems, KGF analogs can be rapidly screened for their biological activity.

The KGF analogs may be further modified to contain additional chemical moieties not normally a part of the peptide. Such derivatized moieties may improve

35   the solubility, absorption, biological half life, and the like of the KGF analog. The moieties may

alternatively eliminate or attenuate any undesirable
side effects of the protein and the like. Moieties
capable of mediating such effects are disclosed, for
example, in *REMINGTON'S PHARMACEUTICAL SCIENCES*, 18th

5   ed., Mack Publishing Co., Easton, PA (1990). Covalent
modifications may be introduced into the molecule by
reacting targeted amino acid residues of the peptide
with an organic derivatizing agent that is capable of
reacting with selected side chains or terminal residues

10 (T.E. Creighton (1983), *PROTEINS: STRUCTURE AND MOLECULE
PROPERTIES*, W.H. Freeman & Co., San Francisco,
pp. 79-86). Polyethylene glycol ("PEG") is one such
chemical moiety which has been used in the preparation
of therapeutic protein products. For some proteins, the

15 attachment of polyethylene glycol has been shown to
protect against proteolysis, Sada, *et al.*(1991), *J.
Fermentation Bioengineering*, 71:137-139 , and methods
for attachment of certain polyethylene glycol moieties
are available. See U.S. Patent No. 4,179,337, Davis *et*

20 *al.*, "Non-Immunogenic Polypeptides," issued December 18,
1979; and U.S. Patent No. 4,002,531, Royer, "Modifying
enzymes with Polyethylene Glycol and Product Produced
Thereby," issued January 11, 1977. For a review, see
Abuchowski *et al.*, in Enzymes as Drugs. (Holcerberg and

25 Roberts, (eds.) pp. 367-383 (1981)). For polyethylene
glycol, a variety of means have been used to attach the
polyethylene glycol molecules to the protein.
Generally, polyethylene glycol molecules are connected
to the protein via a reactive group found on the

30 protein. Amino groups, such as those on lysine residues
or at the N-terminus, are convenient for such
attachment. For example, Royer (U.S. Pat.
No. 4,002,531, above) states that reductive alkylation
was used for attachment of poly thylene glycol molecules

35 to an enzyme. EP 0 539 167, published April 28, 1993,
Wright, "Peg Imidates and Protein Derivates Thereof"

- 20 -

states that peptides and organic compounds with free
amino group(s) are modified with an imidate derivative
of PEG or related water-soluble organic polymers. U.S.
Patent No. 4,904,584, Shaw, issued February 27, 1990,
5   relates to the modification of the number of lysine
residues in proteins for the attachment of polyethylene
glycol molecules via reactive amine groups.

        In yet another embodiment, the present
invention is directed to a single-dose administration
10   unit of a medicinal formulation which can be safely
administered parenterally or orally to treat a disease
in a warm-blooded animal (such as a human). Such
medicinal formulation may be in the form of a
lyophilized or otherwise dehydrated therapeutic or
15   diagnostic which can be reconstituted by the addition of
a physiologically acceptable solvent. The solvent may
be any media such as sterile water, physiological saline
solution, glucose solution or other aqueous
carbohydrates (e.g., polyols such as mannitol, xylitol,
20   glycerol) which is capable of dissolving the dried
composition, is compatible with the selected
administration route and which does not negatively
interfere with the active principle and the
reconstitution stabilizers employed. In a specific
25   embodiment, the present invention is directed to a kit
for producing the single-dose administration unit. The
kit contains both a first container having a dried
protein and a second container having an aqueous
formulation comprising a reconstitution stabilizer. As
30   for the concentration of the protein in the solution,
the solution volume which is charged into each
container, and the capacity of the containers
(interrelated parameters which can be suitably modified,
depending upon the desired concentration of active
35   principl in the end-dosage unit), these may vary within
wide ranges w 11-known to skilled artisans.

KGF analogs according to the invention may be useful as therapeutic and diagnostic agents and as research reagents. Thus the KGF analogs may be used in *in vitro* and/or *in vivo* diagnostic assays to quantify the amount of KGF in a tissue or organ sample or to determine and/or isolate cells which express KGFR (Bottaro *et al.* (1990), *J. Biol. Chem.*, 265:12767-12770; Ron *et al.* (1993), *J. Biol. Chem.*, 268:2984-2988). In assays of tissues or organs there will be less radioactivity from ^{125}I-KGF analog binding to KGFR, as compared to a standardized binding curve of ^{125}I-KGF analog, due to unlabeled native KGF binding to KGFR. Similary, the use of ^{125}I-KGF analog may be used to detect the presence of KGFR in various cell types.

This invention also contemplates the use of a KGF analog in the generation of antibodies made against the peptide, which antibodies also bind to native KGF. In this embodiment, the antibodies are monoclonal or polyclonal in origin and are generated using a KGF analog. The resulting antibodies bind preferentially to native KGF, preferably when that protein is in its native (biologically active) conformation. These antibodies can be used for detection or purification of the KGF.

Moreover, the invention contemplates the use of KGF analogs in the discovery of high affinity or low affinity KGF binding molecules having therapeutical applications, for example, as a way for efficient KGF delivery or as an inhibitor for KGF activity. The thermal stability of the KGF analogs is important to identify such binding molecules in physiological conditions (i.e., at 37°C) since their affinity for KGF could be strongly temperature-dependent and may be unpredictable from the affinity observed at 4°C.

For *in vivo* us s, the KGF analogs may be formulated with additives. Such additives include

buffers, carriers, stabilizers, excipients,
preservatives, tonicity adjusting agents, anti-oxidants
and the like (e.g., viscosity adjusting agents or
extenders). The selection of specific additives will
depend upon the storage form (i.e., liquid or
lyophilized) and the modes of administering the KGF
analog. Suitable formulations, known in the art, can be
found in REMINGTON'S PHARMACEUTICAL SCIENCES (latest
edition), Mack Publishing Company, Easton, PA.

        The KGF analogs may be applied in
therapeutically effective amounts to tissues
specifically characterized by having damage to or
clinically insufficient numbers of non-fibroblast
epithelium cells. Since KGF binds to heparin, it is
likely that heparin, heparin sulfate, heparin-like
glycosaminglycans and heparin-like glycosaminoglycans,
which are present in the extracellular environment may
bind KGF in vivo. It follows that KGF analogs with
reduced heparin binding ability will have enhanced
potency, as more KGF will reach its targeted receptor
and will not be sequestered by heparin and heparin-like
compounds in the extracellular environment. These
analogs will be more useful therapeutically, as lower
dosages of a particular KGF analog will be required per
treatment.

        The KGF analogs may be applied in
therapeutically effective amounts to tissues
specifically characterized by having damage to or
clinically insufficient numbers of non-fibroblast
epithelium cells. Areas in which KGF analogs may be
successfully administered include, but are not limited
to: the stimulation, proliferation and differentiation
of adnexal structures such as hair follicles, sweat
glands, and sebaceous glands in patients with burns and
other partial and full-thickness injuries; accelerat d
r epithelialization of lesions caused by epidermolysis

bullosa, which is a defect in adherence of the epidermis
to the underlying dermis, resulting in frequent open,
painful blisters which can cause severe morbidity;
preventing chemotherapy-induced alopecia and treating
5   male-pattern baldness, or the progressive loss of hair
in men and women; treating gastric and duodenal ulcers;
treating inflammatory bowel diseases, such a Crohn's
disease (affecting primarily the small intestine) and
ulcerative colitis (affecting primarily the large
10   bowel); preventing or reducing gut toxicity in radiation
and chemotherapy treatment regimes through treatment
(e.g., pretreatment and/or posttreatment) to induce a
cytoprotective effect or regeneration or both;
stimulating the production of mucus throughout the
15   gastrointestinal tract; inducing the proliferation and
differentiation of type II pneumocytes, which may help
treat or prevent diseases such as hyaline membrane
disease (i.e., infant respiratory distress syndrome and
bronchopulmonary dysplasia) in premature infants;
20   stimulating the proliferation and differentiation of the
bronchiolar and/or alveolar epithelium with acute or
chronic lung damage or insufficiency due to inhalation
injuries (including high oxygen levels), emphysema, use
of lung damaging chemotherapeutics, ventilator trauma or
25   other lung damaging circumstances; increasing liver
function to treat or prevent hepatic cirrhosis,
fulminant liver failure, damage caused by acute viral
hepatitis and/or toxic insults to the liver; inducing
corneal cell regeneration, for example in the treatment
30   of corneal abrasion; inducing epithelial cell
regeneration to treat progressive gum disease; inducing
regeneration of tympanic epithelial cells to treat ear
drum damage and treating or preventing the onset of
diabetes mellitus or as an adjunct in the setting of
35   islet c 11 transplantation.

        A patient in need of proliferation of non-
fibroblast epithelial cells will be administered an
effective amount of a KGF analog. An "effective amount"
is that amount of KGF analog required to elicit the
5   desired response in the patient being treated and will,
thus, generally be determined by the attending
physician. Factors influencing the amount of KGF analog
administered will include the age and general condition
of the patient, the disease being treated, etc. Typical
10   dosages will range from 0.001 mg/kg body weight to 500
mg/kg body weight.

        The KGF analog may be safely administered
parenterally (e.g., via IV, IT, IM, SC, or IP routes),
orally or topically to warm-blooded animals (such as
15   humans). The KGF analog may be used once or
administered repeatedly, depending on the disease and
condition of the patient. In some cases, the KGF analog
may be administered as an adjunct to other therapy and
also with other pharmaceutical preparations.
20   The following examples are included to more
fully illustrate the present invention. It is
understood that modifications can be made in the
procedures set forth, without departing from the spirit
of the invention.
25

                          EXAMPLES

        Standard methods for many of the procedures
described in the following examples, or suitable
30   alternative procedures, are provided in widely
recognized manuals of molecular biology such as, for
example, *Molecular Cloning*, Second Edition, Sambrook *et
al.*, Cold Spring Harbor Laboratory Press (1987) and
Current Protocols in Molecular Biology, Ausabel *et al.*,
35   Greene Publishing Associates/Wiley Interscience, New
York (1990).

EXAMPLE 1: Preparation of DNA Coding for KGF and KGF Analogs

The cloning of the full-length human KGF gene
5   (encoding a polypeptide with the sequence of native KGF)
was carried out both by polymerase chain reaction (PCR)
of RNA from an animal cell and by PCR of chemically
synthesized (*E. coli* optimized codon) oligonucleotides
("OLIGOs"). Both procedures are described below:
10   PCR amplification using RNA isolated from
cells known to produce the polypeptide was performed.
Initially, cells from a human fibroblast cell line
AG1523A (obtained from Human Genetic Mutant Cell Culture
Repository Institute For Medical Research, Camden, New
15 Jersey) were disrupted with guanidium thiocyanate,
followed by extraction (according to the method of
Chomyzinski *et al.* (1987), *Anal. Biochem.*, 172:156).
Using a standard reverse transcriptase protocol for
total RNA, the KGF cDNA was generated. PCR (PCR#1)
20 amplification of the KGF gene was carried out using the
KGF cDNA as template and primers OLIGO#1 and OLIGO#2
that encode DNA sequences immediately 5' and 3' of the
KGF gene [Model 9600 thermocycler (Perkin-Elmer Cetus,
Norwalk, CT); 28 cycles; each cycle consisting of one
25 minute at 94°C for denaturation, two minutes at 60°C for
annealing, and three minutes at 72°C for elongation]. A
small aliquot of the PCR#1 product was then used as
template for a second KGF PCR (PCR#2) amplification
identical to the cycle conditions described above except
30 for a 50°C annealing temperature. For expression
cloning of the KGF gene, nested PCR primers were used to
create convenient restriction sites at both ends of the
KGF gene. OLIGO#3 and OLIGO#4 were used to modify the
KGF DNA product from PCR#2 to include *MluI* and *BamHI*
35 restriction sites at the 5' and 3' ends of the gene,
respectively [PCR#3; 30 cycles; each cycle consisting of

one minute at 94°C for denaturation, two minutes at 60°C
for annealing, and three minutes at 72°C for
elongation]. This DNA was subsequently cut with *MluI*
and *BamHI*, phenol extracted, and ethanol precipitated.

5   It was then resuspended and ligated (using T4 ligase)
into a pCFM1156 plasmid (Figure 2A) that contained a
"RSH" signal sequence to make construct RSH-KGF
(Figure 3).

The ligation products were transformed

10   (according to the method of Hanahan (1983), *J. Mol.
Biol.*, <u>166</u>:557) into *E. coli* strain FM5 (ATCC: 53911)
and plated onto LB+kanamycin at 28°C. Several
transformants were selected and grown in small liquid
cultures containing 20 µg/mL kanamycin. The RSH-KGF

15   plasmid was isolated from the cells of each culture and
DNA sequenced. Because of an internal *NdeI* site in the
KGF gene, it was not possible to directly clone the
native gene sequence into the desired expression vector
with the bracketed restriction sites of *NdeI* and *BamHI*.

20   This was accomplished as a three-way ligation. Plasmid
RSH-KGF was cut with the unique restriction sites of
BsmI and *SstI*, and a ~3 kbp DNA fragment (containing the
3' end of the KGF gene) was isolated following
electrophoresis through a 1% agarose gel. A PCR (PCR#4)

25   was carried out as described for PCR#3 except for the
substitution of OLIGO#5 for OLIGO#3. The PCR DNA
product was then cut with *NdeI* and *BsmI* and a 311 bp DNA
fragment was isolated following electrophoresis through
a 4% agarose gel. The third piece of the ligation is a

30   1.8 kbp DNA fragment of pCFM1156 cut with *NdeI* and *SstI*
which was isolated following electrophoresis through a
1% agarose gel. Following ligation (T4 ligase),
transformation, kanamycin selection and DNA sequencing
as described above, a clone was picked containing the

35   construct in Figure 4 and the plasmid d signated KGF.
Because of an internal ribosomal binding site that

produced truncated products, the KGF DNA sequence
between the unique *KpnI* and *EcoRI* sites was replaced
with chemically synthesized OLIGOs (OLIGO#6 through
OLIGO#11) to minimize the use of the internal start site
5   (Figure 5).

    OLIGO#1 (SEQ ID NO:7):   5'-CAATGACCTAGGAGTAACAATCAAC-3'
    OLIGO#2 (SEQ ID NO:8):   5'-AAAACAAACATAAATGCACAAGTCCA-3'
    OLIGO#3 (SEQ ID NO:9):   5'-ACAACGCGTGCAATGACATGACTCCA-3'
10 OLIGO#4 (SEQ ID NO:10):
              5'-ACAGGATCCTATTAAGTTATTGCCATAGGAA-3'
    OLIGO#5 (SEQ ID NO:11):
              5'-ACACATATGTGCAATGACATGACTCCA-3'
    OLIGO#6 (SEQ ID NO:12):
15   5'-CTGCGTATCGACAAACGCGGCAAAGTCAAGGGCACCC-3'
    OLIGO#7 (SEQ ID NO:13):
              5'-AAGAGATGAAAAACAACTACAATATTATGGAAATCCGTACTGTT-3'
    OLIGO#8 (SEQ ID NO:14):
              5'-GCTGTTGGTATCGTTGCAATCAAAGGTGTTGAATCTG-3'
20 OLIGO#9 (SEQ ID NO:15):
              5'-TCTTGGGTGCCCTTGACTTTGCCGCGTTTGTCGATACGCAGGTAC-3'
    OLIGO#10 (SEQ ID NO:16):
              5'-ACAGCAACAGTACGGATTTCCATAATATTGTAGTTGTTTTTCATC-3'
    OLIGO#11 (SEQ ID NO:17):
25   5'-AATTCAGATTCAACACCTTTGATTGCAACGATACCA-3'

         The OLIGOs were phosphorylated with T4
polynucleotide kinase and then heat denatured. The
single-stranded (ss) OLIGOs were then allowed to form a
30 ds DNA fragment by allowing the temperature to slowly
decrease to room temperature. T4 ligase was then used
to covalently link both the internal OLIGO sticky-ends
and the whole ds OLIGO fragment to the KGF plasmid cut
with *KpnI* and *EcoRI*. The new plasmid was designated
35 KGF(dsd).

A completely *E. coli* codon-optimized KGF gene
was constructed by PCR amplification of chemically
synthesized OLIGOs #12 through 24.

5   OLIGO#12 (SEQ ID NO:18):   5'-AGTTTTGATCTAGAAGGAGG-3'

       OLIGO#13 (SEQ ID NO:19):   5'-TCAAAACTGGATCCTATTAA-3'

       OLIGO#14 (SEQ ID NO:20):

              5'-AGTTTTGATCTAGAAGGAGGAATAACATATGTGCAACGACATG-

              ACTCCGGAACAGATGGCTACCAACGTTAACTGCTCCAGCCCGGAACGT-3'

10   OLIGO#15 (SEQ ID NO:21):

              5'-CACACCCGTAGCTACGACTACATGGAAGGTGGTGACATCCGT-

              GTTCGTCGTCTGTTCTGCCGTACCCAGTGGTACCTGCGTATCGACAAA-3'

       OLIGO#16 (SEQ ID NO:22):

              5'-CGTGGTAAAGTTAAAGGTACCCAGGAAATGAAAAACAACTACAACATC-

15   ATGGAAATCCGTACTGTTGCTGTTGGTATCGTTGCAATCAAA-3'

       OLIGO#17 (SEQ ID NO:23):

              5'-GGTGTTGAATCTGAATTCTACCTGGCAATGAACAAAGAAGGTAAACT-

              GTACGCAAAAAAAGAATGCAACGAAGACTGCAACTTCAAAGAA-3'

       OLIGO#18 (SEQ ID NO:24):

20   5'-CTGATCCTGGAAAACCACTACAACACCTACGCATCTGCTAAATGGAC-

              CCACAACGGTGGTGAAATGTTCGTTGCTCTGAACCAGAAAGGT-3'

       OLIGO#19 (SEQ ID NO:25):

              5'-ATCCCGGTTCGTGGTAAAAAAACCAAAAAAGAACAGAAAACCGCTC-

              ACTTCCTGCCGATGGCAATCACTTAATAGGATCCAGTTTTGA-3'

25   OLIGO#20 (SEQ ID NO:26):5'-TACGGGTGTGACGTTCCGGG-3'

       OLIGO#21 (SEQ ID NO:27):5'-CTTTACCACGTTTGTCGATA-3'

       OLIGO#22 (SEQ ID NO:28):5'-ATTCAACACCTTTGATTGCA-3'

       OLIGO#23 (SEQ ID NO:29):5'-CCAGGATCAGTTCTTTGAAG-3'

       OLIGO#24 (SEQ ID NO:30):5'-GAACCGGGATACCTTTCTGG-3'

30

OLIGOs #12 through 24 were designed so that
the entire DNA sequence encoding native KGF was
represented by OLIGOs from either the "Watson" or the
"Crick" strand and upon PCR amplification would produce
35 the desired double-stranded DNA sequence (Figure 6)
[PCR#5, Model 9600 thermocycler, Perkin-Elmer Cetus]; 21

cycles, each cycle consisting of 31 seconds at 94°C for denaturation, 31 seconds at 50°C for annealing, and 31 seconds at 73°C for elongation; following the 21 cycles the PCR was finished with a final elongation step of 7

5 minutes]. After PCR amplification, the DNA fragment was cut with *XbaI* and *BamHI* and the 521 bp fragment ligated into the expression plasmid pCFM1156 cut with the same enzymes. PCR#5 utilized the outside primers (100 pmoles/100 µl rxn) OLIGO#12 and OLIGO#13 and 1 µl/100 µl

10 rxn of a KGF template derived by ligation (by T4 ligase) of OLIGO #14 through OLIGO#19 (OLIGO#15 through OLIGO#18 were phosphorylated with T4 polynucleotide kinase) using OLIGO#20 through OLIGO#24 as band-aid oligos (Jayaraman *et al*. (1992), *Biotechniques*, 12:392) for the

15 ligation. The final construct was designated KGF (codon optimized).

All of the KGF analogs described herein are composed in part from DNA sequences found in KGF(dsd) or KGF(codon optimized), or a combination of the two. The

20 sequences are further modified by the insertion into convenient restriction sites of DNA sequences that encode the particular KGF analog amino acids made utilizing one or more of the above-described techniques for DNA fragment synthesis. Any of the analogs can be

25 generated in their entirety by the above described techniques. However, as a part of the general OLIGO design optimized *E. coli* codons were used where appropriate, although the presence of *E. coli* optimized codons in part or *in toto* of any of the genes where

30 examined did not significantly increase the yield of protein that could be obtained from cultured bacterial cells. Figures 7 to 10 and 17 to 26 set forth by convenient xample particular KGF analog nucleotide and amino acid sequence constructions: R(144)Q (Figure 7);

35 C(1,15)S/R(144)E (Figure 8); C(1,15)S/R(144)Q

(Figure 9); ΔN23/R(144)Q (Figure 10); ΔN23/N(137)E
(Figure 17); ΔN23/K(139)E (Figure 18); ΔN23/K(139)Q
(Figure 19); ΔN23/R(144)A (Figure 20); ΔN23/R(144)L
(Figure 21); ΔN23/K(147)E (Figure 22); ΔN23/K(147)Q
5   (Figure 23); ΔN23/K(153)E (Figure 24); ΔN23/K(153)Q;
(Figure 25) and ΔN23/Q(152)E/K(153)E (Figure 26). All
the KGF analog constructions described herein were DNA
sequence confirmed.

10   EXAMPLE 2: Production in *E. coli*

        Three different expression plasmids were
utilized in the cloning of the KGF analog genes. They
were pCFM1156 (ATCC# 69702), pCFM1656 (ATCC# 69576), and
15   pCFM3102 (Figures 2A, 2B and 2C, respectively). The
plasmid p3102 can be derived from the plasmid pCFM1656
by making a series of site-directed base changes with
PCR overlapping oligo mutagenesis. Starting with the
BglII site (pCFM1656 plasmid bp # 180) immediately 5' to
20   the plasmid replication promoter, P_{copB}, and proceeding
toward the plasmid replication genes, the base pair
changes are as follows:

pCFM1656 bp #	bp in pCFM1656	bp changed to in pCFM3102
# 204	T/A	C/G
# 428	A/T	G/C
# 509	G/C	A/T
# 617	- -	insert two G/C bp
# 677	G/C	T/A
# 978	T/A	C/G
# 992	G/C	A/T
# 1002	A/T	C/G
# 1005	C/G	T/A
# 1026	A/T	T/A
# 1045	C/G	T/A
# 1176	G/C	T/A
# 1464	G/C	T/A
# 2026	G/C	bp deletion
# 2186	C/G	T/A
# 2479	A/T	T/A
# 2498-2501	AGTG	GTCA
	TCAC	CAGT
# 2641-2647	TCCGAGC	bp deletion
	AGGCTCG	
# 3441	G/C	A/T
# 3452	G/C	A/T
# 3649	A/T	T/A
# 4556	--	insert bps

```
(SEQ ID NO:39) 5'-GAGCTCACTAGTGTCGACCTGCAG-3'
(SEQ ID NO:40) 3'-CTCGAGTGATCACAGCTGGACGTC-5'
```

As seen above, pCFM1156, pCFM1656 and pCFM3102 are very similar to each other and contain many of the same restriction sites. The plasmids were chosen by convenience, and the vector DNA components can be easily exchanged for purposes of new constructs. The host used for all cloning was *E. coli* strain FM5 (ATCC: 53911) and the transformations were carried out (according to the method of Hanahan (1983), *supra*) or by electroelution with a Gene Pulser™ transfection apparatus (BioRad Laboratories, Inc., Hercules, CA) according to the manufacturer's instructions.

Initially, a small, freshly cultured inoculum of the desired recombinant *E. coli* clone harboring the

desired construct on one of the three pCFM vectors was
started by transferring 0.1 mL of a frozen glycerol
stock of the appropriate strain into a 2 L flask
containing 500 mL of Luria broth. The culture was
5   shaken at 30˚C for 16 hours, after which the culture was
transferred to a 15 L fermentor containing 8 L of
sterile batch medium (Tsai, et al. (1987), *J. Industrial
Microbiol.*, 2:181-187).

Feed batch fermentation starts with the
10 feeding of Feed # 1 medium (Tsai, et al. (1987),
supra). When the OD600 reached 35, expression of the
desired KGF analog was induced by rapidly raising the
culture temperature to 37˚C for two hours then up to
42˚C to denature the CI repressor. The addition of Feed
15 1 was discontinued in favor of Feed 2, the addition rate
of which was initiated at 300 mL/hr. Feed 2 comprised
175 g/L trypticase-peptone, 87.5 g/L yeast extract, and
260 g/L glucose. After one hour at 42˚C, the culture
temperature was decreased to 36˚C, where this
20 temperature was then maintained for another 6 hours.

The fermentation was then halted and the cells
were harvested by centrifugation into plastic bags
placed within 1 L centrifuge bottles. The cells were
pelleted by centrifugation at 400 rpm for 60 minutes,
25 after which the supernatants were removed and the cell
paste frozen at -90˚C.

Following expression of the various KGF
analogs, in *E. coli*, native KGF, R(144)Q,
C(1,15)S/R(144)E, C(1,15)S/R(144)Q and ΔN23/R(144)Q
30 proteins were purified using the following procedure.
Cell paste from a high cell density fermentation was
suspended at 4˚C in 0.2 M NaCl, 20 mM NaPO$_4$, pH 7.5 as a
10-20% solution (w ight per volume) using a suitable
high shear mixer. The suspended cells were then lysed
35 by passing the solution through a homogenizer (APV
Gaulin, Inc., Everett, MA) three times. The outflowing

homogenate was cooled to 4-8°C by using a suitable heat exchanger. Debris was then removed by centrifuging the lysate in a J-6B™ centrifuge (Beckman Instruments, Inc., Brea, CA) equipped with a JS 4.2 rotor at 4,200 rpm for
5 30-60 min. at 4°C. Supernatants were then carefully decanted and loaded onto a previously prepared 450 mL (5 cm x 23 cm) column of S-Sepharose Fast Flow™ resin (Pharmacia, Piscataway, NJ) equilibrated with 0.2 M NaCl, 20 mM NaPO$_4$, pH 7.5 at 4°C. Next, the column was
10 washed with five column volumes (2250 mL) of 0.4 M NaCl, 20 mM NaPO$_4$, pH 7.5 at 4°C. The desired protein was eluted by washing the column with 5 L of 0.5 M NaCl, 20 mM NaPO$_4$, pH 7.5. Then, 50 mL fractions were collected and the A$_{280}$ of the effluent was continuously monitored.
15 Fractions identified by A$_{280}$ as containing eluted material were then analyzed by SDS-PAGE through 14% gels to confirm the presence of the desired polypeptide.

Those fractions containing proteins of interest were then pooled, followed by the addition of
20 an equal volume of distilled water. The diluted sample was then loaded onto a previously prepared 450 mL (5 cm x 23 cm) column of S-Sepharose Fast Flow equilibrated with 0.4 M NaCl, 20 mM NaPO$_4$, pH 6.8 at 4°C. The column was washed with 2250 mL of 0.4 M NaCl, 20 mM NaPO$_4$, pH
25 6.8 and the protein eluted using a 20 column volume linear gradient ranging from 0.4 M NaCl, 20 mM NaPO$_4$, pH 6.8 to 0.6 M NaCl, 20 mM NaPO$_4$, pH 6.8. Again, 50 mL fractions were collected under constant A$_{280}$ monitoring of the effluent. Those fractions containing the protein
30 (determined by 14% SDS-PAGE) were then pooled, followed by concentration through a YM-10 membrane (10,000 molecular weight cutoff) in a 350cc stirring cell (Amicon, Inc. Mayberry, MA) to a volume of 30-40 mL.

The concentrate was then loaded onto a
35 previously generated 1,300 mL (4.4 cm x 85 cm) column of Superdex-75™ r sin (Pharmacia) quilibrated in column

- 34 -

buffer comprising 1X PBS (Dulbecco's Phosphate Buffered
Saline, "D-PBS", calcium and magnesium-free) or 0.15 M
NaCl, 20 mM NaPO$_4$, pH 7.0. After allowing the sample to
run into the column, the protein was eluted from the gel
5   filtration matrix using column buffer. Thereafter, 10
mL fractions were recovered and those containing the
analog (determined by 14% SDS-PAGE) were pooled.
Typically, the protein concentration was about 5-10
mg/mL in the resultant pool. All of the above
10   procedures were performed at 4-8°C, unless otherwise
specified.

Analysis

15   Analysis was conducted on *E. coli*-derived,
native KGF; R(144)Q; C(1,15)S/R(144)E; C(1,15)S/R(144)Q
and ΔN23/R(144)Q.

Conformational Stability

20
The polypeptides were compared by their
storage stability, thermal unfolding transition
temperatures (T$_m$), and stability in a broad range of pH
conditions.

25   The ability of native KGF, R(144)Q,
C(1,15)S/R(144)Q, C(1,15)S/R(144)E and ΔN23/R(144)Q
to prevent aggregation at elevated temperatures was also
examined. Samples containing 0.5 mg/mL of protein were
prepared in D-PBS. 0.5 mL of each sample was aliquoted
30   into 3 cc type-1 glass vials. The vials were sealed
with rubber stoppers and 13 mm flip-off aluminum seals
were crimped on. These vials were then placed in a 37°C
incubator. At predetermined time intervals, vials were
withdrawn and analyzed for the loss of soluble protein.
35   Visible precipitat s were removed by centrifuging 250 μL
of each sample through a 0.22 μm Spin-X filter unit

(Costar, Cambridge, MA). Soluble protein in the filtered solutions was subsequently analyzed by size exclusion HPLC. The amount of soluble protein was determined by integrating the HPLC peak area and

5 plotting the result as a function of incubation time at 37°C. The results are shown in Figure 11.

The half-lives for the loss of soluble, monomeric protein were then estimated from these kinetic curves. Table 1 shows the half-life for remaining

10 soluble KGF upon storage at 37°C for these proteins.

Table 1
Half-life for the Loss of Soluble, Monomeric Proteins

Protein	t1/2 (day)
native KGF	0.6
R(144)Q	4.1
C(1,15)S/R(144)Q	13.3
ΔN23/R(144)Q	22.3
C(1,15)S/R(144)E	38.0

15

As seen in Table 1, above, and Figure 11, the native KGF aggregated the most rapidly, with a half-life of 0.6 days. R(144)Q increased the half-life to 4.1 days. C(1,15)S/R(144)Q, ΔN23/R(144)Q and

20 C(1,15)S/R(144)E showed substantial increases in the solubility half-life to 13.3, 22.3 and 38 days, respectively.

Thermal Unfolding

25

Thermal unfolding was monitored by circular dichroism (CD) at 230 nm using a J-720™ spectropolarimeter (Jasco, Inc., Easton, MD) equipped with a PTC-343 Peltier-type temprature control system.

30 For CD analysis, separate samples containing 0.1 mg/mL

of the polypeptide to be analyzed were prepared in D-PBS
(Life Technologies, Inc., Grand Island, NY). For each
sample, about 2.5 mL was loaded into a 10 mm path length
rectangular Suprasil™ quartz (Heraeus Quarzschmelze,
5   GmbH, Hanau, Germany) fluorescent cell (Hellma Cells,
Inc., Jamaica, NY). The cell was then placed into the
Peltier-type temperature control system in the
spectropolarimeter. Thermal unfolding was carried out
at a rate of 50°C/hr. Changes in ellipticity were
10 monitored at 230 nm to indicate unfolding. The T_m of
each sample was estimated by identifying a temperature
at which 50% of protein molecules in the solution were
unfolded (*Biophysical Chemistry*, Cantor and Schimmel
(eds), W.H. Freeman and Co. San Francisco (1980)). The
15 estimated T_m for each of the three proteins is listed in
Table 2.

Table 2
Estimated Melting Temperatures

20

Protein	T_m (°C)
native KGF	54.0
R(144)Q	61.5
C(1,15)S/R(144)Q	62.5
ΔN23/R(144)Q	63.0
C(1,15)S/R(144)E	63.5

As the results show, R(144)Q has a greater than 7°C
increase in the Tm as compared with native KGF. The
substitution of R(144)Q to C(1,15)S/R(144)Q or ΔN23 adds
25 at least another 1°C increase in T_m and more than 8°C as
compared with native KGF. Moreover, the
C(1,15)S/R(144)E is greater than 9°C more stable than
native KGF. Therefore, switching a positiv ly charged
residue (Arg) at amino acid position 144 to a neutrally

or negatively charged residue substantially stabilized
the polypeptide.

                                pH
5

        The acid stabilities of C(1,15)S/R(144)Q and
C(1,15)S/R(144)E were also compared to that of native
KGF, by adjusting D-PBS to different pH values by adding
concentrated HCl or NaOH. Approximately 2.35 mL of
10   D-PBS at different pH values was mixed with 100 µL of
2.45 mg/mL KGF protein in a quartz cell. These samples
were thermally unfolded at a rate of 50°C/hr and
monitored by CD at 230 nm. Figure 12 shows the T_m as a
function of pH for native KGF, C(1,15)S/R(144)Q and
15   C(1,15)S/R(144)E. In the pH range tested, the
C(1,15)S/R(144)Q and C(1,15)S/R(144)E always have a
higher T_m than the native KGF.

                    In vitro Biological Activity
20

        In vitro mitogenic activity of R(144)Q,
ΔN23/R(144)Q, C(1,15)S/R(144)Q and C(1,15)S/R(144)E was
also determined as a function of protein concentration
and the half-maximal concentrations by measurement of
25   [^{3}H]-thymidine uptake by Balb/MK cells (according to the
methods of Rubin _et al._ (1989), _supra_).
        Generally, the concentrations of each of the
KGF analogs relative to a known standard native KGF was
determined using an _in vitro_ biological assay. Each KGF
30   analog was then diluted and assayed for biological
activity using a Balb/MK mitogenic assay. The samples
were first diluted in a bioassay medium consisting of
50% customer-made Eagle's MEM, 50% customer-made F12,
5 µG/mL transferrin, 5 ng/ml sodium selenite, 0.0005%
35   HSA and 0.005% Tween 20. KGF samples were then added
into Falcon primeria 96-well plates seeded with Balb/MK

- 38 -

cells. Incorporation of [^{3}H]-Thymidine during DNA
synthesis was measured and converted to input native KGF
concentration by comparison to a native KGF standard
curve. The results are presented in Figures 13 to 16.
5 As seen in Figures 13 to 16, each of the KGF analogs has
mitogenic activity.

While the present invention has been described
above both generally and in terms of preferred
embodiments, it is understood that other variations and
10 modifications will occur to those skilled in the art in
light of the description above.

WHAT IS CLAIMED IS:

1.   A polypeptide analog of native KGF comprising
a charge-change by the deletion or substitution of one
5   or more of amino acid residues 41-154 of Figure 2 (amino
acids 72-185 of SEQ ID NO:2).

2.   The polypeptide analog according to Claim 1
wherein the deleted or substituted amino acids are
10   selected from the group consisting of Arg^{41}, Gln^{43},
Lys^{55}, Lys^{95}, Lys^{128}, Asn^{137}, Gln^{138}, Lys^{139}, Arg^{144},
Lys^{147}, Gln^{152}, Lys^{153} and Thr^{154}.

3.   The polypeptide analog according to Claim 1
15   selected from the group consisting of R(144)Q,
C(1,15)S/R(144)Q, C(1,15)S/R(144)E and ΔN23/R(144)Q.

4.   A pharmaceutical formulation comprising a
therapeutically effective amount of a polypeptide analog
20   of KGF according to Claim 1 and a pharmaceutically
acceptable carrier.

5.   A pharmaceutical formulation comprising a
therapeutically effective amount of a lyophilized
25   polypeptide analog of KGF according to Claim 1.

6.   The pharmaceutical formulation of Claim 4
further comprising a pharmaceutically acceptable
carrier.
30

7.   A nucleic acid molecule selected from the
group consisting of DNA and RNA wherein the nucleic acid
mol cule encodes a polypeptide analog of native KGF
comprising a charge-change by the deletion or
35   substitution of one or mor of amino acid residues 41-
154 of Figure 2 (amino acids 72-185 of SEQ ID NO:2).

8.   The nucleic acid molecule according to Claim 7
wherein the deleted or substituted amino acids are
selected from the group consisting of Arg^{41}, Gln^{43},
5   Lys^{55}, Lys^{95}, Lys^{128}, Asn^{137}, Gln^{138}, Lys^{139}, Arg^{144},
Lys^{147}, Gln^{152}, Lys^{153} and Thr^{154}.

9.   The nucleic acid molecule according to Claim 7
wherein the polypeptide analog is selected from the
10   group consisting of R(144)Q, C(1,15)S/R(144)Q,
C(1,15)S/R(144)E and ΔN23/R(144)Q.

10.   A biologically functional plasmid or viral
vector comprising a nucleic acid molecule according to
15   Claim 7.

11.   A procaryotic or eucaryotic host cell stably
transfected or transformed with a biologically
functional vector according to Claim 8.
20

12.   A procaryotic host cell according to Claim 11
that is *E. coli*.

13.   A eucaryotic host cell according to Claim 11
25   that is a mammalian cell.

14.   A eucaryotic host cell according to Claim 12
that is a Chinese hamster ovary cell.

30   15.   A process for the production of a polypeptide
analog of KGF, the process comprising growing under
suitable nutrient conditions a procaryotic or eucaryotic
host cell stably transformed with a nucleic acid
molecul according to Claim 7, in a manner allowing
35   expression of the encoded polypeptide analog, and
isolating the polypeptide analog so produced.

16.   A method of stimulating the production of
non-fibroblast epithelial cells comprising contacting
such cells with an effective amount of a polypeptide
5   analog of a KGF according to Claim 1.

Figure 1

human KGF (+ signal sequenc)

```
                              |-------OLIGO#1---------|
                              5'CAATGACCTAGGAGTAACAATCAAC 3'
5'CAATCTACAATTCACAGATAGGAAGAGGTCAATGACCTAGGAGTAACAATCAACTCAAGA-
 ---------+---------+---------+---------+---------+---------+ 60

-TTCATTTTCATTATGTTATTCATGAACACCCGGAGCACTACACTATAATGCACAAATGGA-
 ---------+---------+---------+---------+---------+---------+ 120
                                                     M   H   K   W   I

-TACTGACATGGATCCTGCCAACTTTGCTCTACAGATCATGCTTTCACATTATCTGTCTAG-
 ---------+---------+---------+---------+---------+---------+ 180
   L   T   W   I   L   P   T   L   L   Y   R   S   C   F   H   I   I   C   L   V

-TGGGTACTATATCTTTAGCTTGCAATGACATGACTCCAGAGCAAATGGCTACAAATGTGA-
 ---------+---------+---------+---------+---------+---------+ 240
     G   T   I   S   L   A   C   N   D   M   T   P   E   Q   M   A   T   N   V   N

-ACTGTTCCAGCCCTGAGCGACACACAAGAAGTTATGATTACATGGAAGGAGGGGATATAA-
 ---------+---------+---------+---------+---------+---------+ 300
     C   S   S   P   E   R   H   T   R   S   Y   D   Y   M   E   G   G   D   I   R

-GAGTGAGAAGACTCTTCTGTCGAACACAGTGGTACCTGAGGATCGATAAAAGAGGCAAAG-
 ---------+---------+---------+---------+---------+---------+ 360
     V   R   R   L   F   C   R   T   Q   W   Y   L   R   I   D   K   R   G   K   V

-TAAAAGGGACCCAAGAGATGAAGAATAATTACAATATCATGGAAATCAGGACAGTGGCAG-
 ---------+---------+---------+---------+---------+---------+ 420
     K   G   T   Q   E   M   K   N   N   Y   N   I   M   E   I   R   T   V   A   V

-TTGGAATTGTGGCAATCAAAGGGGTGGAAAGTGAATTCTATCTTGCAATGAACAAGGAAG-
 ---------+---------+---------+---------+---------+---------+ 480
     G   I   V   A   I   K   G   V   E   S   E   F   Y   L   A   M   N   K   E   G

-GAAAACTCTATGCAAAGAAAGAATGCAATGAAGATTGTAACTTCAAAGAACTAATTCTGG-
 ---------+---------+---------+---------+---------+---------+ 540
     K   L   Y   A   K   K   E   C   N   E   D   C   N   F   K   E   L   I   L   E

-AAAACCATTACAACACATATGCATCAGCTAAATGGACACACAACGGAGGGGAAATGTTTG-
 ---------+---------+---------+---------+---------+---------+ 600
     N   H   Y   N   T   Y   A   S   A   K   W   T   H   N   G   G   E   M   F   V

-TTGCCTTAAATCAAAAGGGGATTCCTGTAAGAGGAAAAAAAACGAAGAAAGAACAAAAAA-
 ---------+---------+---------+---------+---------+---------+ 660
     A   L   N   Q   K   G   I   P   V   R   G   K   K   T   K   K   E   Q   K   T

-CAGCCCACTTTCTTCCTATGGCAATAACTTAATTGCATATGGTATATAAAGAACCCAGTT
 ---------+---------+---------+---------+---------+---------+ 720
     A   H   F   L   P   M   A   I   T   *
```

Figure 1
(c ntinued)

```
-CCAGCAGGGAGATTTCTTTAAGTGGACTGTTTTCTTTCTTCTCAAAATTTTCTTTCCTTT
 ---------+---------+---------+---------+---------+---------+ 780

-TATTTTTTAGTAATCAAGAAAGGCTGGAAAAACTACTGAAAAACTGATCAAGCTGGACTT
 ---------+---------+---------+---------+---------+---------+ 840
                                                  3'ACCTGAA-
                                                   |------

-GTGCATTTATGTTTGTTTTAAG 3'
 ---------+---------+-- 862
-CACGTAAATACAAACAAAA 5'
 ---OLIGO#2-------|
```

Figure 2A

Figure 2B

Figure 2C
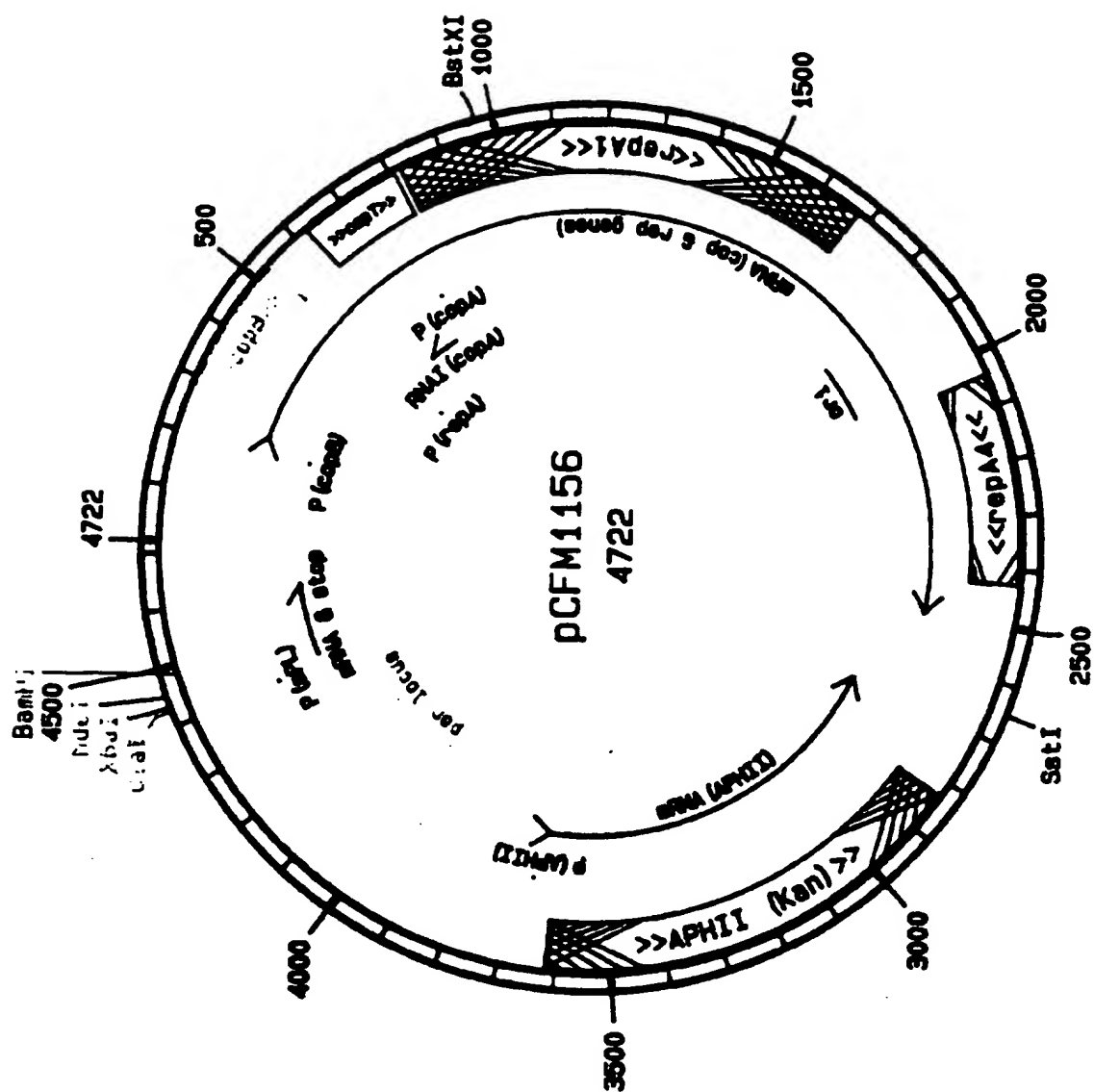
Figure 3
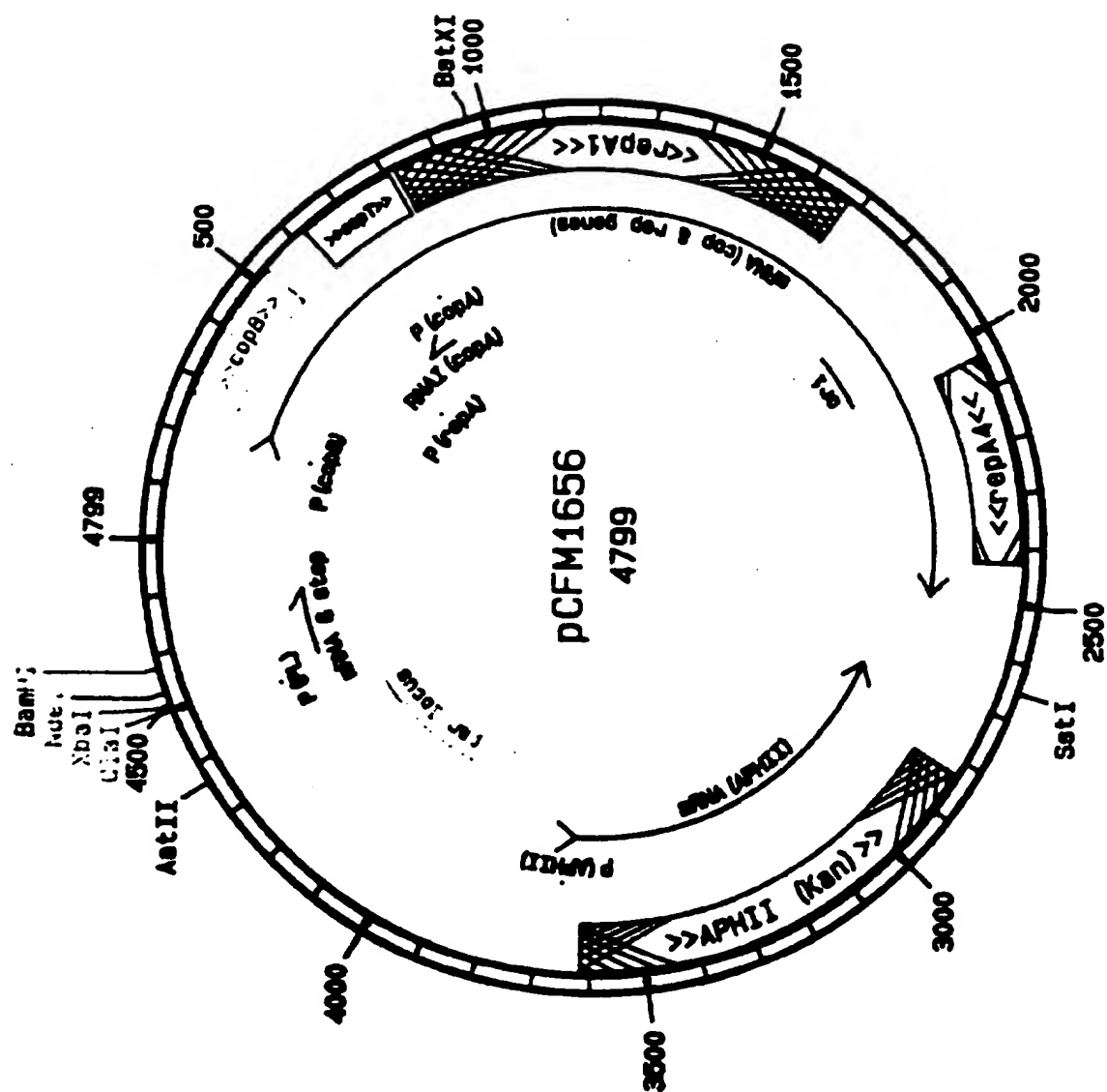
RSH-KGF
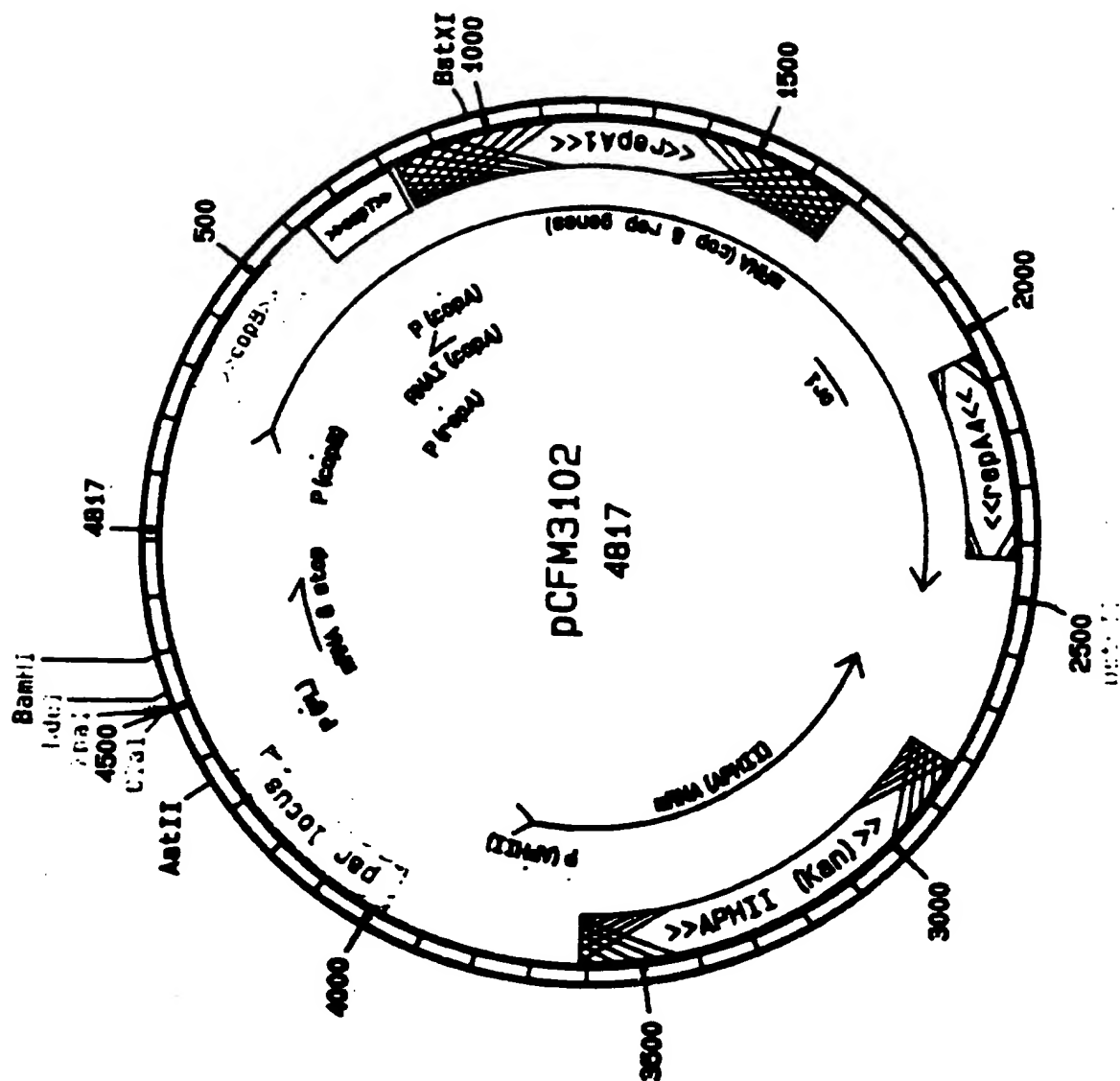
```
plasmid DNA              ClaI        XbaI                    NdeI
sequence         5'-ATCGATTTGATTCTAGAAGGAGGAATAACATATGAAAAAG-
                                                            M  K  K

             RSH signal sequence                           MluI
 -CGCGCACGTGCTATCGCCATTGCTGTGGCTCTGGCAGGTTTCGCAACTAGTGCACA-3'
   R  A  R  A  I  A  I  A  V  A  L  A  G  F  A  T  S  A  H  A  -


    MluI
 5'CGCGTGCAATGACATGACTCCAGAGCAAATGGCTACAAATGTGAACTGTTCCAGCCCTGA-
   ---------+---------+---------+---------+---------+---------+  60
     -  C  N  D  M  T  P  E  Q  M  A  T  N  V  N  C  S  S  P  E

 -GCGACACACAAGAAGTTATGATTACATGGAAGGAGGGGATATAAGAGTGAGAAGACTCTT-
   ---------+---------+---------+---------+---------+---------+  120
    R  H  T  R  S  Y  D  Y  M  E  G  G  D  I  R  V  R  R  L  F

                   KpnI        ClaI
 -CTGTCGAACACAGTGGTACCTGAGGATCGATAAAAGAGGCAAAGTAAAAGGGACCCAAGA-
   ---------+---------+---------+---------+---------+---------+  180
    C  R  T  Q  W  Y  L  R  I  D  K  R  G  K  V  K  G  T  Q  E

 -GATGAAGAATAATTACAATATCATGGAAATCAGGACAGTGGCAGTTGGAATTGTGGCAAT-
   ---------+---------+---------+---------+---------+---------+  240
    M  K  N  N  Y  N  I  M  E  I  R  T  V  A  V  G  I  V  A  I

                   EcoRI
 -CAAAGGGGTGGAAAGTGAATTCTATCTTGCAATGAACAAGGAAGGAAAACTCTATGCAAA-
   ---------+---------+---------+---------+---------+---------+  300
    K  G  V  E  S  E  F  Y  L  A  M  N  K  E  G  K  L  Y  A  K

             BsmI
 -GAAAGAATGCAATGAAGATTGTAACTTCAAAGAACTAATTCTGGAAAACCATTACAACAC-
   ---------+---------+---------+---------+---------+---------+  360
    K  E  C  N  E  D  C  N  F  K  E  L  I  L  E  N  H  Y  N  T

    NdeI
 -ATATGCATCAGCTAAATGGACACACAACGGAGGGGAAATGTTTGTTGCCTTAAATCAAAA-
   ---------+---------+---------+---------+---------+---------+  420
    Y  A  S  A  K  W  T  H  N  G  G  E  M  F  V  A  L  N  Q  K

 -GGGGATTCCTGTAAGAGGAAAAAAAACGAAGAAAGAACAAAAAAACAGCCCACTTTCTTCC-
   ---------+---------+---------+---------+---------+---------+  480
    G  I  P  V  R  G  K  K  T  K  K  E  Q  K  T  A  H  F  L  P

             BamHI
 -TATGGCAATAACTTAATAG 3'                 -plasmid DNA
   ---------+---------+--- 503            -s qu nc
    M  A  I  T      *
```

Figure 4

KGF

NdeI
5'TATGTGCAATGACATGACTCCAGAGCAAATGGCTACAAATGTGAACTGTTCCAGCCCTGA-
---------+---------+---------+---------+---------+---------+ 60
    M   C   N   D   M   T   P   E   Q   M   A   T   N   V   N   C   S   S   P   E

-GCGACACACAAGAAGTTATGATTACATGGAAGGAGGGGATATAAGAGTGAGAAGACTCTT-
---------+---------+---------+---------+---------+---------+ 120
    R   H   T   R   S   Y   D   Y   M   E   G   G   D   I   R   V   R   R   L   F

                KpnI   ClaI
-CTGTCGAACACAGTGGTACCTGAGGATCGATAAAAGAGGCAAAGTAAAAGGGACCCAAGA-
---------+---------+---------+---------+---------+---------+ 180
    C   R   T   Q   W   Y   L   R   I   D   K   R   G   K   V   K   G   T   Q   E

-GATGAAGAATAATTACAATATCATGGAAATCAGGACAGTGGCAGTTGGAATTGTGGCAAT-
---------+---------+---------+---------+---------+---------+ 240
    M   K   N   N   Y   N   I   M   E   I   R   T   V   A   V   G   I   V   A   I

                EcoRI
-CAAAGGGGTGGAAAGTGAATTCTATCTTGCAATGAACAAGGAAGGAAAACTCTATGCAAA-
---------+---------+---------+---------+---------+---------+ 300
    K   G   V   E   S   E   F   Y   L   A   M   N   K   E   G   K   L   Y   A   K

            BsmI
-GAAAGAATGCAATGAAGATTGTAACTTCAAAGAACTAATTCTGGAAAACCATTACAACAC-
---------+---------+---------+---------+---------+---------+ 360
    K   E   C   N   E   D   C   N   F   K   E   L   I   L   E   N   H   Y   N   T

    NdeI
-ATATGCATCAGCTAAATGGACACACAACGGAGGGGAAATGTTTGTTGCCTTAAATCAAAA-
---------+---------+---------+---------+---------+---------+ 420
    Y   A   S   A   K   W   T   H   N   G   G   E   M   F   V   A   L   N   Q   K

-GGGGATTCCTGTAAGAGGAAAAAAAACGAAGAAAGAACAAAAAACAGCCCACTTTCTTCC-
---------+---------+---------+---------+---------+---------+ 480
    G   I   P   V   R   G   K   K   T   K   K   E   Q   K   T   A   H   F   L   P

                BamHI
-TATGGCAATAACTTAATAG 3'
---------+---------+--- 503
    M   A   I   T   *

Figure 5

subsitution of KpnI to EcoRI sequenc to make KGF(dsd)

```
KpnI
    |-----------OLIGO#6-----------------||------OLIGO#7---------
5'    CTGCGTATCGACAAACGCGGCAAAGTCAAGGGCACCCAAGAGATGAAAAACAACTACAAT-
3'CATGGACGCATAGCTGTTTGCGCCGTTTCAGTTCCCGTGGGTTCTCTACTTTTTGTTGATGTTA-
    |---------------OLIGO#9--------------------||--OLIGO#10--------
  - Y   L   R   I   D   K   R   G   K   V   K   G   T   Q   E   M   K   N   N   Y   N  -
```

```
                                                                   EcoRI
    --------------------------||---------OLIGO#8-------------------|
5'-ATTATGGAAATCCGTACTGTTGCTGTTGGTATCGTTGCAATCAAAGGTGTTGAATCTG        3'
3'-TAATACCTTTAGGCATGACAACGACAACCATAGCAACGTTAGTTTCCACAACTTAGACTTAA 5'
    -------------------------||----OLIGO#11---------------------|
    I   M   E   I   R   T   V   A   V   G   I   V   A   I   K   G   V   E   S   E   F  -
```

FIGURE 6

KGF (codon optimized)

```
          XbaI
     |-----OLIGO#12-----|
 5'AGTTTTGATCTAGAAGGAGG 3'
     |----------------------------OLIGO#14---------------------------
 5'AGTTTTGATCTAGAAGGAGGAATAACATATGTCAACGACATGACTCCGGAACAGATGGCT-


      --------------------------------||--------------OLIGO#15--------
     -ACCAACGTTAACTGCTCCAGCCCGGAACGTCACACCCGTAGCTACGACTACATGGAAGGTG-
                    3' GGGCCTTGCAGTGTGGGCAT 5'
                       |-----OLIGO#20-----|


      ------------------------------OLIGO#15----------------------------||-
     -GTGACATCCGTGTTCGTCGTCTGTTCTGCCGTACCCAGTGGTACCTGCGTATCGACAAACG-
                                                  3' ATAGCTGTTTGC-
                                                  |-OLIGO#21--


      ------------------------------OLIGO#16-----------------------------
     -TGGTAAAGTTAAAGGTACCCAGGAAATGAAAAACAACTACAACATCATGGAAATCCGTACT-
     -ACCATTTC 5'
      -------|


      --------------------------------||--------------OLIGO#17-----------
     -GTTGCTGTTGGTATCGTTGCAATCAAAGGTGTTGAATCTGAATTCTACCTGGCAATGAACA-
                    3' ACGTTAGTTTCCACAACTTA 5'
                       |-----OLIGO#22-----|


      ------------------OLIGO#17---------------------------------||----
     -AAGAAGGTAAACTGTACGCAAAAAAAGAATGCAACGAAGACTGCAACTTCAAAGAACTGAT-
                                                  3'GAAGTTTCTTGACTA-
                                                  |---OLIGO#23---


      ----------------------OLIGO#18-----------------------------------
     -CCTGGAAAACCACTACAACACCTACGCATCTGCTAAATGGACCCACAACGGTGGTGAAATG-
     -GGACC 5'
      ----|


      ----------------------------||--------------------OLIGO#19---------
     -TTCGTTGCTCTGAACCAGAAAGGTATCCCGGTTCGTGGTAAAAAAACCAAAAAAGAACAGA-
                    3' GGTCTTTCCATAGGGCCAAG 5'
                       |-----OLIGO#24-----|


      -------OLIGO#19---------------------------------------|
     -AAACCGCTCACTTCCTGCCGATGGCAATCACTTAATAGGATCCAGTTTTGA 3'
                                 3' AATTATCCTAGGTCAAAACT 5'
                                 |----OLIGO#13-----|
                                      BamHI
```

Figure 7

KGF R(144)Q

```
5'ATGTGCAATGATATGACTCCTGAACAAATGGCTACCAATGTCAACTGTTCCTCTCCGGAG-
  ---------+---------+---------+---------+---------+---------+ 60
   M  C  N  D  M  T  P  E  Q  M  A  T  N  V  N  C  S  S  P  E

 -CGCCACACCCGGAGTTACGATTACATGGAAGGTGGGGATATTCGCGTACGTCGTCTGTTC-
  ---------+---------+---------+---------+---------+---------+ 120
   R  H  T  R  S  Y  D  Y  M  E  G  G  D  I  R  V  R  R  L  F

 -TGCCGTACCCAGTGGTACCTGCGTATCGACAAACGCGGCAAAGTCAAGGGCACCCAAGAG-
  ---------+---------+---------+---------+---------+---------+ 180
   C  R  T  Q  W  Y  L  R  I  D  K  R  G  K  V  K  G  T  Q  E

 -ATGAAAAACAACTACAATATTATGGAAATCCGTACTGTTGCTGTTGGTATCGTTGCAATC-
  ---------+---------+---------+---------+---------+---------+ 240
   M  K  N  N  Y  N  I  M  E  I  R  T  V  A  V  G  I  V  A  I

 -AAAGGTGTTGAATCTGAATTCTATCTTGCAATGAACAAGGAAGGAAAACTCTATGCAAAG-
  ---------+---------+---------+---------+---------+---------+ 300
   K  G  V  E  S  E  F  Y  L  A  M  N  K  E  G  K  L  Y  A  K

 -AAAGAATGCAATGAAGATTGTAACTTCAAAGAACTAATTCTGGAAAACCATTACAACACA-
  ---------+---------+---------+---------+---------+---------+ 360
   K  E  C  N  E  D  C  N  F  K  E  L  I  L  E  N  H  Y  N  T

 -TATGCATCTGCTAAATGGACCCACAACGGTGGTGAAATGTTCGTTGCTCTGAACCAGAAA-
  ---------+---------+---------+---------+---------+---------+ 420
   Y  A  S  A  K  W  T  H  N  G  G  E  M  F  V  A  L  N  Q  K

 -GGTATCCCTGTTCAAGGTAAGAAAACCAAGAAAGAACAGAAAACCGCTCACTTCCTGCCG-
  ---------+---------+---------+---------+---------+---------+ 480
   G  I  P  V  Q  G  K  K  T  K  K  E  Q  K  T  A  H  F  L  P

 -ATGGCAATCACTTAA 3'
  ---------+----- 495
   M  A  I  T  *
```

Figur 8

KGF C(1,15)S/R(144)E

```
5'ATGTCTAATGATATGACTCCGGAACAGATGGCTACCAACGTTAACTCCTCCTCCCCGGAA-
  ---------+---------+---------+---------+---------+---------+ 60
   M   S   N   D   M   T   P   E   Q   M   A   T   N   V   N   S   S   S   P   E

-CGTCACACGCGTTCCTACGACTACATGGAAGGTGGTGACATCCGCGTACGTCGTCTGTTC-
  ---------+---------+---------+---------+---------+---------+ 120
   R   H   T   R   S   Y   D   Y   M   E   G   G   D   I   R   V   R   R   L   F

-TGCCGTACCCAGTGGTACCTGCGTATCGACAAACGCGGCAAAGTCAAGGGCACCCAAGAG-
  ---------+---------+---------+---------+---------+---------+ 180
   C   R   T   Q   W   Y   L   R   I   D   K   R   G   K   V   K   G   T   Q   E

-ATGAAAAACAACTACAATATTATGGAAATCCGTACTGTTGCTGTTGGTATCGTTGCAATC-
  ---------+---------+---------+---------+---------+---------+ 240
   M   K   N   N   Y   N   I   M   E   I   R   T   V   A   V   G   I   V   A   I

-AAAGGTGTTGAATCTGAATTCTATCTTGCAATGAACAAGGAAGGAAAACTCTATGCAAAG-
  ---------+---------+---------+---------+---------+---------+ 300
   K   G   V   E   S   E   F   Y   L   A   M   N   K   E   G   K   L   Y   A   K

-AAAGAATGCAATGAAGATTGTAACTTCAAAGAACTAATTCTGGAAAACCATTACAACACA-
  ---------+---------+---------+---------+---------+---------+ 360
   K   E   C   N   E   D   C   N   F   K   E   L   I   L   E   N   H   Y   N   T

-TATGCATCTGCTAAATGGACCCACAACGGTGGTGAAATGTTCGTTGCTCTGAACCAGAAA-
  ---------+---------+---------+---------+---------+---------+ 420
   Y   A   S   A   K   W   T   H   N   G   G   E   M   F   V   A   L   N   Q   K

-GGTATCCCTGTTGAAGGTAAGAAAACCAAGAAAGAACAGAAAACCGCTCACTTCCTGCCG-
  ---------+---------+---------+---------+---------+---------+ 480
   G   I   P   V   E   G   K   K   T   K   K   E   Q   K   T   A   H   F   L   P

-ATGGCAATCACTTAA 3'
  ---------+----- 495
   M   A   I   T   *
```

Figure 9

KGF C(1,15)S/R144Q

```
5'ATGTCTAATGATATGACTCCGGAACAGATGGCTACCAACGTTAACTCCTCCTCCCCGGAA-
  ---------+---------+---------+---------+---------+---------+ 60
   M  S  N  D  M  T  P  E  Q  M  A  T  N  V  N  S  S  S  P  E

 -CGTCACACGCGTTCCTACGACTACATGGAAGGTGGTGACATCCGCGTACGTCGTCTGTTC-
  ---------+---------+---------+---------+---------+---------+ 120
   R  H  T  R  S  Y  D  Y  M  E  G  G  D  I  R  V  R  R  L  F

 -TGCCGTACCCAGTGGTACCTGCGTATCGACAAACGCGGCAAAGTCAAGGGCACCCAAGAG-
  ---------+---------+---------+---------+---------+---------+ 180
   C  R  T  Q  W  Y  L  R  I  D  K  R  G  K  V  K  G  T  Q  E

 -ATGAAAAACAACTACAATATTATGGAAATCCGTACTGTTGCTGTTGGTATCGTTGCAATC-
  ---------+---------+---------+---------+---------+---------+ 240
   M  K  N  N  Y  N  I  M  E  I  R  T  V  A  V  G  I  V  A  I

 -AAAGGTGTTGAATCTGAATTCTATCTTGCAATGAACAAGGAAGGAAAACTCTATGCAAAG-
  ---------+---------+---------+---------+---------+---------+ 300
   K  G  V  E  S  E  F  Y  L  A  M  N  K  E  G  K  L  Y  A  K

 -AAAGAATGCAATGAAGATTGTAACTTCAAAGAACTAATTCTGGAAAACCATTACAACACA-
  ---------+---------+---------+---------+---------+---------+ 360
   K  E  C  N  E  D  C  N  F  K  E  L  I  L  E  N  H  Y  N  T

 -TATGCATCTGCTAAATGGACCCACAACGGTGGTGAAATGTTCGTTGCTCTGAACCAGAAA-
  ---------+---------+---------+---------+---------+---------+ 420
   Y  A  S  A  K  W  T  H  N  G  G  E  M  F  V  A  L  N  Q  K

 -GGTATCCCTGTTCAAGGTAAGAAAACCAAGAAAGAACAGAAAACCGCTCACTTCCTGCCG-
  ---------+---------+---------+---------+---------+---------+ 480
   G  I  P  V  Q  G  K  K  T  K  K  E  Q  K  T  A  H  F  L  P

 -ATGGCAATCACTTAA 3'
  ---------+----- 495
   M  A  I  T  *
```

Figure 10

KGF ΔN23/R(144)Q

```
5'ATGTCCTACGACTACATGGAAGGTGGTGACATCCGCGTACGTCGTCTGTTCTGCCGTACC-
  ---------+---------+---------+---------+---------+---------+ 60
   M  S  Y  D  Y  M  E  G  G  D  I  R  V  R  R  L  F  C  R  T

 -CAGTGGTACCTGCGTATCGACAAACGCGGCAAAGTCAAGGGCACCCAAGAGATGAAAAAC-
  ---------+---------+---------+---------+---------+---------+ 120
   Q  W  Y  L  R  I  D  K  R  G  K  V  K  G  T  Q  E  M  K  N

 -AACTACAATATTATGGAAATCCGTACTGTTGCTGTTGGTATCGTTGCAATCAAAGGTGTT-
  ---------+---------+---------+---------+---------+---------+ 180
   N  Y  N  I  M  E  I  R  T  V  A  V  G  I  V  A  I  K  G  V

 -GAATCTGAATTCTATCTTGCAATGAACAAGGAAGGAAAACTCTATGCAAAGAAAGAATGC-
  ---------+---------+---------+---------+---------+---------+ 240
   E  S  E  F  Y  L  A  M  N  K  E  G  K  L  Y  A  K  K  E  C-

 -AATGAAGATTGTAACTTCAAAGAACTAATTCTGGAAAACCATTACAACACATATGCATCT-
  ---------+---------+---------+---------+---------+---------+ 300
   N  E  D  C  N  F  K  E  L  I  L  E  N  H  Y  N  T  Y  A  S

 -GCTAAATGGACCCACAACGGTGGTGAAATGTTCGTTGCTCTGAACCAGAAAGGTATCCCT-
  ---------+---------+---------+---------+---------+---------+ 360
   A  K  W  T  H  N  G  G  E  M  F  V  A  L  N  Q  K  G  I  P

 -GTTCAAGGTAAGAAAACCAAGAAAGAACAGAAAACCGCTCACTTCCTGCCGATGGCAATC-
  ---------+---------+---------+---------+---------+---------+ 420
   V  Q  G  K  K  T  K  K  E  Q  K  T  A  H  F  L  P  M  A  I

 -ACTTAA 3'
  ------ 426
   T  *
```

Figure 11

Figure 12

Thermal Unfolding of KGF Analogs

Figure 13

Figure 14

wt vs. delta N23/R144Q
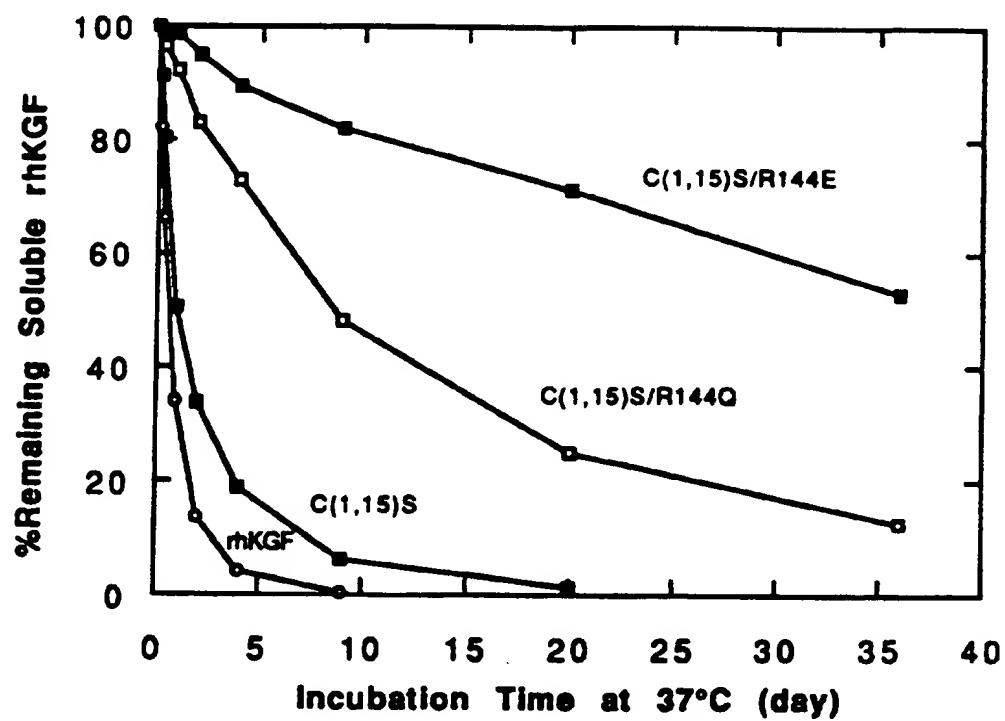
Figure 15

Figure 16

wt vs. C1,15S/R144E

Figure 17

ΔN23/N(137)E

```
5'-ATGTCCTACGACTACATGGAAGGTGGTGACATCCGCGTACGTCGTCTGTTCTGCCGTACC-
   +---------+---------+---------+---------+---------+---------
    M  S  Y  D  Y  M  E  G  G  D  I  R  V  R  R  L  F  C  R  T  -

  -CAGTGGTACCTGCGTATCGACAAACGCGGCAAAGTCAAGGGCACCCAAGAGATGAAAAAC-
   +---------+---------+---------+---------+---------+---------
    Q  W  Y  L  R  I  D  K  R  G  K  V  K  G  T  Q  E  M  K  N  -

  -AACTACAATATTATGGAAATCCGTACTGTTGCTGTTGGTATCGTTGCAATCAAAGGTGTT-
   +---------+---------+---------+---------+---------+---------
    N  Y  N  I  M  E  I  R  T  V  A  V  G  I  V  A  I  K  G  V  -

  -GAATCTGAATTCTACCTGGCAATGAACAAAGAAGGTAAACTGTACGCAAAAAAAGAATGC-
   +---------+---------+---------+---------+---------+---------
    E  S  E  F  Y  L  A  M  N  K  E  G  K  L  Y  A  K  K  E  C  -

  -AACGAAGACTGCAACTTCAAAGAACTGATCCTGGAAAACCACTACAACACCTACGCATCT-
   +---------+---------+---------+---------+---------+---------
    N  E  D  C  N  F  K  E  L  I  L  E  N  H  Y  N  T  Y  A  S  -

  -GCTAAATGGACCCACAACGGTGGTGAAATGTTCGTTGCTCTGGAACAGAAAGGTATCCCT-
   +---------+---------+---------+---------+---------+---------
    A  K  W  T  H  N  G  G  E  M  F  V  A  L  E  Q  K  G  I  P  -

  -GTTCGTGGTAAGAAAACCAAGAAAGAACAGAAAACCGCTCACTTCCTGCCGATGGCAATC-
   +---------+---------+---------+---------+---------+---------
    V  R  G  K  K  T  K  K  E  Q  K  T  A  H  F  L  P  M  A  I  -

  -ACTTAA-3'
   +-----
    T  *
```
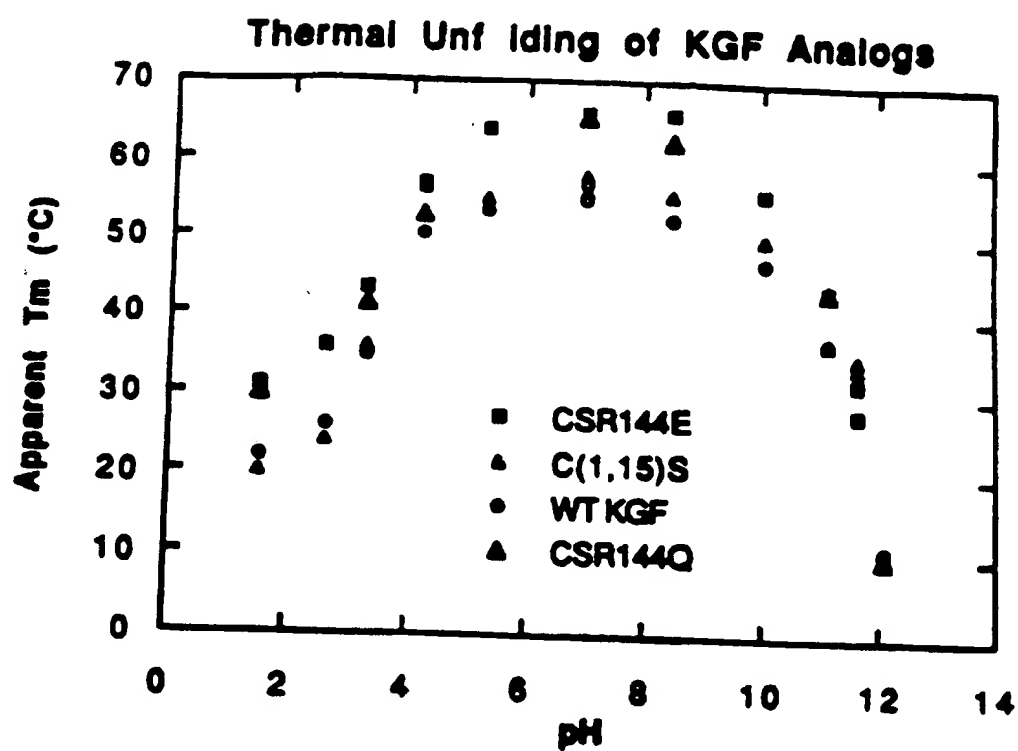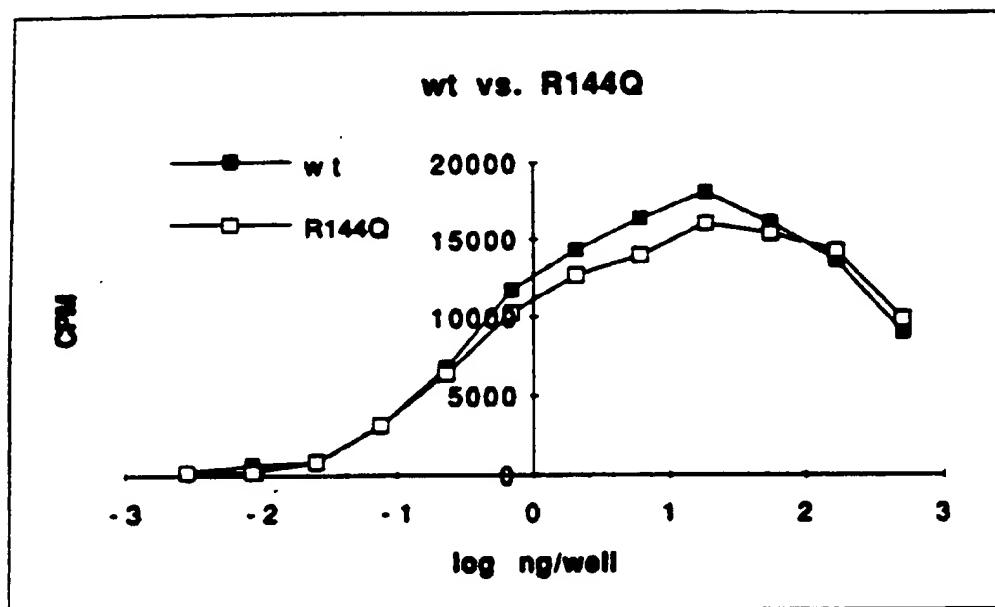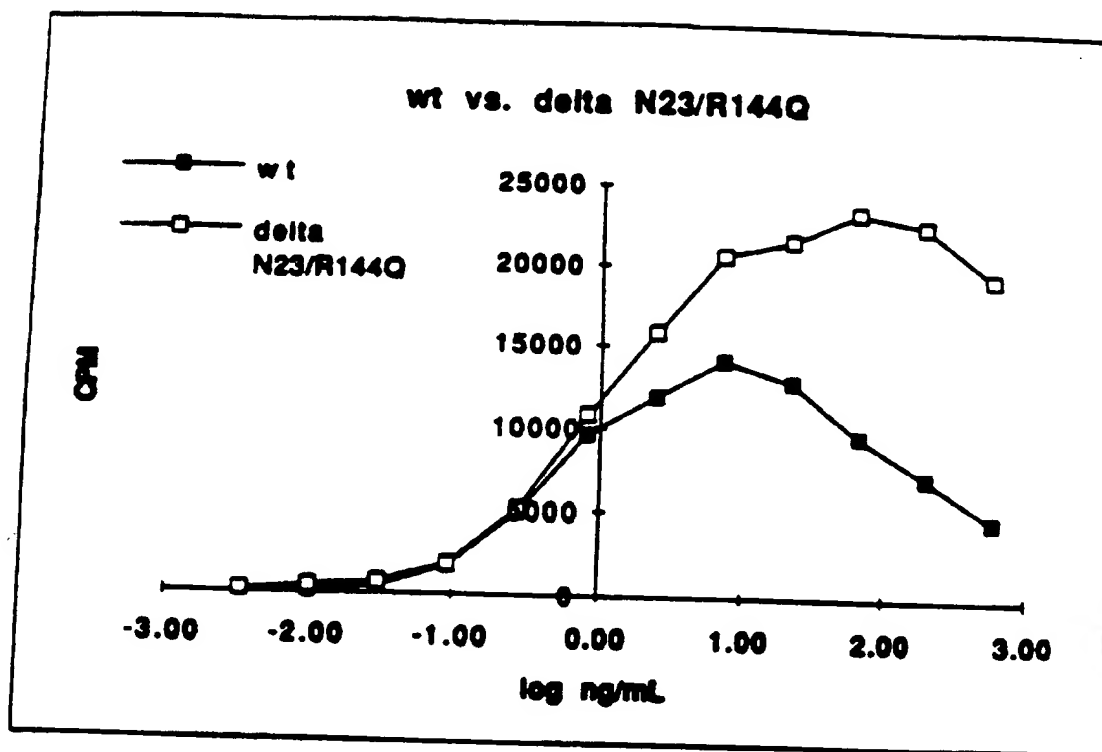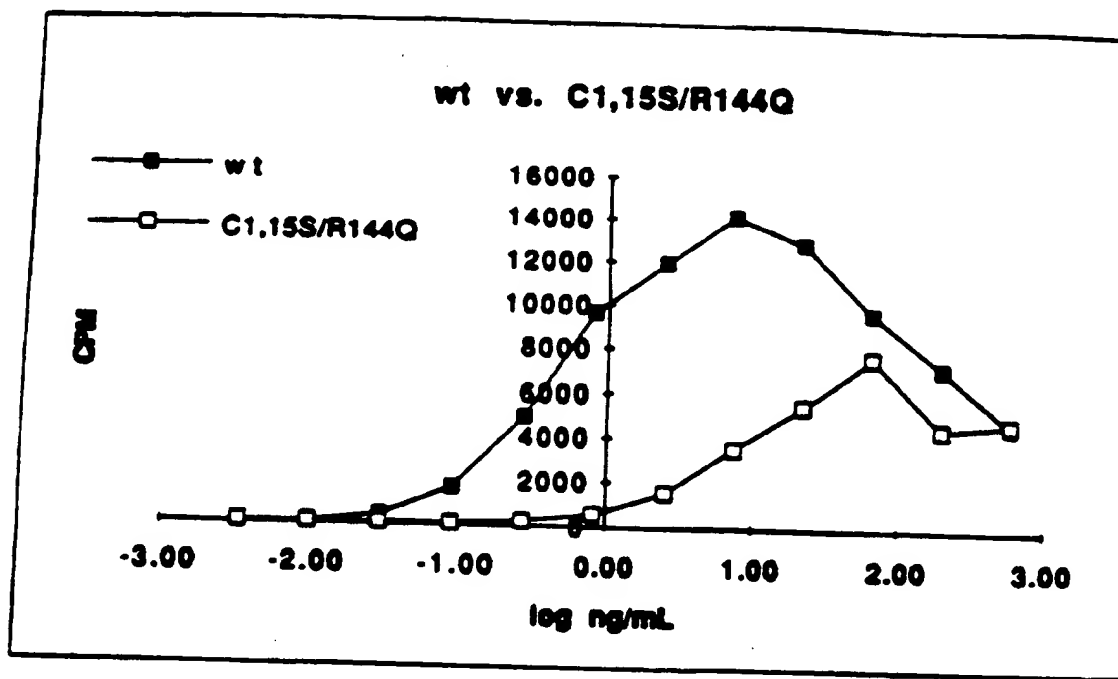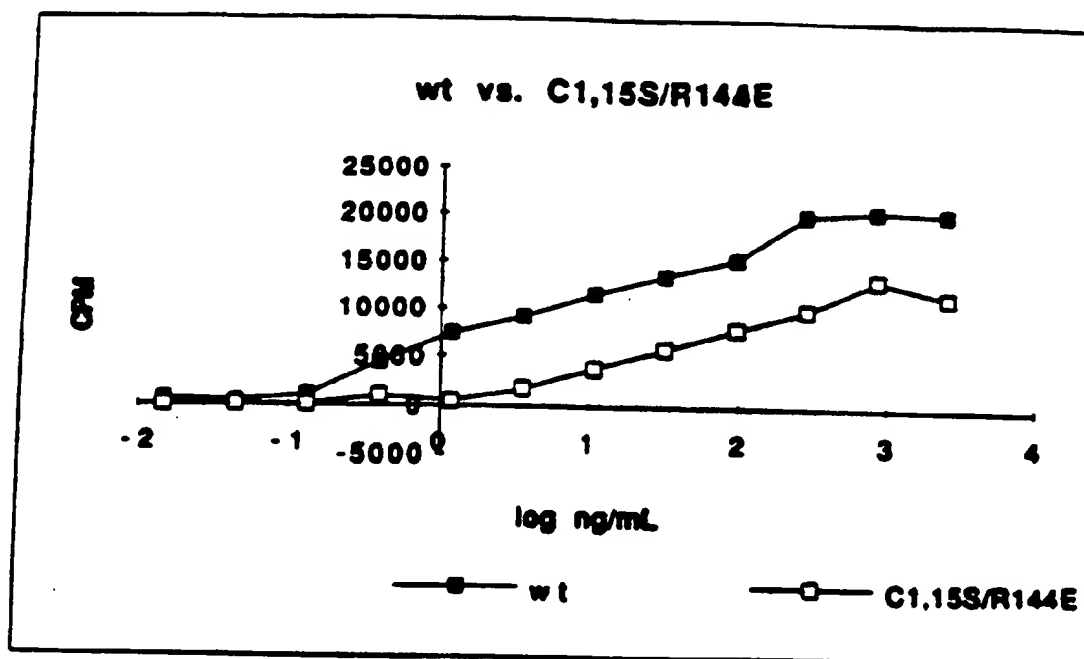
Figure 18

ΔN23/K(139)E

```
5'-ATGTCCTACGACTACATGGAAGGTGGTGACATCCGCGTACGTCGTCTGTTCTGCCGTACC-
   +---------+---------+---------+---------+---------+---------
    M   S   Y   D   Y   M   E   G   G   D   I   R   V   R   R   L   F   C   R   T   -

  -CAGTGGTACCTGCGTATCGACAAACGCGGCAAAGTCAAGGGCACCCAAGAGATGAAAAAC
   +---------+---------+---------+---------+---------+---------
    Q   W   Y   L   R   I   D   K   R   G   K   V   K   G   T   Q   E   M   K   N   -

  -AACTACAATATTATGGAAATCCGTACTGTTGCTGTTGGTATCGTTGCAATCAAAGGTGTT-
   +---------+---------+---------+---------+---------+---------
    N   Y   N   I   M   E   I   R   T   V   A   V   G   I   V   A   I   K   G   V   -

  -GAATCTGAATTCTACCTGGCAATGAACAAAGAAGGTAAACTGTACGCAAAAAAAGAATGC-
   +---------+---------+---------+---------+---------+---------
    E   S   E   F   Y   L   A   M   N   K   E   G   K   L   Y   A   K   K   E   C   -

  -AACGAAGACTGCAACTTCAAAGAACTGATCCTGGAAAACCACTACAACACCTACGCATCT
   +---------+---------+---------+---------+---------+---------
    N   E   D   C   N   F   K   E   L   I   L   E   N   H   Y   N   T   Y   A   S   -

  -GCTAAATGGACCCACAACGGTGGTGAAATGTTCGTTGCTCTGAACCAGGAAGGTATCCCT-
   +---------+---------+---------+---------+---------+---------
    A   K   W   T   H   N   G   G   E   M   F   V   A   L   N   Q   E   G   I   P   -

  -GTTCGTGGTAAGAAAACCAAGAAAGAACAGAAAACCGCTCACTTCCTGCCGATGGCAATC-
   +---------+---------+---------+---------+---------+---------
    V   R   G   K   K   T   K   K   E   Q   K   T   A   H   F   L   P   M   A   I   -

  -ACTTAA-3'
   +-----
    T   *
```

Figure 19

ΔN23/K(139)Q

```
5'-ATGTCCTACGACTACATGGAAGGTGGTGACATCCGCGTACGTCGTCTGTTCTGCCGTACC-
   +---------+---------+---------+---------+---------+---------
    M  S  Y  D  Y  M  E  G  G  D  I  R  V  R  R  L  F  C  R  T  -

  -CAGTGGTACCTGCGTATCGACAAACGCGGCAAAGTCAAGGGCACCCAAGAGATGAAAAAC-
   +---------+---------+---------+---------+---------+---------
    Q  W  Y  L  R  I  D  K  R  G  K  V  K  G  T  Q  E  M  K  N  -

  -AACTACAATATTATGGAAATCCGTACTGTTGCTGTTGGTATCGTTGCAATCAAAGGTGTT-
   +---------+---------+---------+---------+---------+---------
    N  Y  N  I  M  E  I  R  T  V  A  V  G  I  V  A  I  K  G  V  -

  -GAATCTGAATTCTACCTGGCAATGAACAAAGAAGGTAAACTGTACGCAAAAAAAGAATGC-
   +---------+---------+---------+---------+---------+---------
    E  S  E  F  Y  L  A  M  N  K  E  G  K  L  Y  A  K  K  E  C  -

  -AACGAAGACTGCAACTTCAAAGAACTGATCCTGGAAAACCACTACAACACCTACGCATCT-
   +---------+---------+---------+---------+---------+---------
    N  E  D  C  N  F  K  E  L  I  L  E  N  H  Y  N  T  Y  A  S  -

  -GCTAAATGGACCCACAACGGTGGTGAAATGTTCGTTGCTCTGAACCAGCAAGGTATCCCT-
   +---------+---------+---------+---------+---------+---------
    A  K  W  T  H  N  G  G  E  M  F  V  A  L  N  Q  Q  G  I  P  -

  -GTTCGTGGTAAGAAAACCAAGAAAGAACAGAAAACCGCTCACTTCCTGCCGATGGCAATC-
   +---------+---------+---------+---------+---------+---------
    V  R  G  K  K  T  K  K  E  Q  K  T  A  H  F  L  P  M  A  I  -

  -ACTTAA-3'
   +-----
    T  *
```

Figure 20

ΔN23/R(144)A

```
5'-ATGTCCTACGACTACATGGAAGGTGGTGACATCCGCGTACGTCGTCTGTTCTGCCGTACC-
   +---------+---------+---------+---------+---------+---------
    M  S  Y  D  Y  M  E  G  G  D  I  R  V  R  R  L  F  C  R  T  -

  -CAGTGGTACCTGCGTATCGACAAACGCGGCAAAGTCAAGGGCACCCAAGAGATGAAAAAC-
   +---------+---------+---------+---------+---------+---------
    Q  W  Y  L  R  I  D  K  R  G  K  V  K  G  T  Q  E  M  K  N  -

  -AACTACAATATTATGGAAATCCGTACTGTTGCTGTTGGTATCGTTGCAATCAAAGGTGTT-
   +---------+---------+---------+---------+---------+---------
    N  Y  N  I  M  E  I  R  T  V  A  V  G  I  V  A  I  K  G  V  -

  -GAATCTGAATTCTACCTGGCAATGAACAAAGAAGGTAAACTGTACGCAAAAAAGAATGC-
   +---------+---------+---------+---------+---------+---------
    E  S  E  F  Y  L  A  M  N  K  E  G  K  L  Y  A  K  K  E  C  -

  -AACGAAGACTGCAACTTCAAAGAACTGATCCTGGAAAACCACTACAACACCTACGCATCT-
   +---------+---------+---------+---------+---------+---------
    N  E  D  C  N  F  K  E  L  I  L  E  N  H  Y  N  T  Y  A  S  -

  -GCTAAATGGACCCACAACGGTGGTGAAATGTTCGTTGCTCTGAACCAGAAAGGTATCCCT-
   +---------+---------+---------+---------+---------+---------
    A  K  W  T  H  N  G  G  E  M  F  V  A  L  N  Q  K  G  I  P  -

  -GTTGCTGGTAAGAAAACCAAGAAAGAACAGAAAACCGCTCACTTCCTGCCGATGGCAATC-
   +---------+---------+---------+---------+---------+---------
    V  A  G  K  K  T  K  K  E  Q  K  T  A  H  F  L  P  M  A  I  -

  -ACTTAA-3'
   +-----
    T  *
```

Figure 21

ΔN23/R(144)L

```
5'-ATGTCCTACGACTACATGGAAGGTGGTGACATCCGCGTACGTCGTCTGTTCTGCCGTACC-
   +---------+---------+---------+---------+---------+---------
   M  S  Y  D  Y  M  E  G  G  D  I  R  V  R  R  L  F  C  R  T -

-CAGTGGTACCTGCGTATCGACAAACGCGGCAAAGTCAAGGGCACCCAAGAGATGAAAAAC-
 +---------+---------+---------+---------+---------+---------
 Q  W  Y  L  R  I  D  K  R  G  K  V  K  G  T  Q  E  M  K  N -

-AACTACAATATTATGGAAATCCGTACTGTTGCTGTTGGTATCGTTGCAATCAAAGGTGTT-
 +---------+---------+---------+---------+---------+---------
 N  Y  N  I  M  E  I  R  T  V  A  V  G  I  V  A  I  K  G  V -

-GAATCTGAATTCTACCTGGCAATGAACAAAGAAGGTAAACTGTACGCAAAAAAAGAATGC-
 +---------+---------+---------+---------+---------+---------
 E  S  E  F  Y  L  A  M  N  K  E  G  K  L  Y  A  K  K  E  C -

-AACGAAGACTGCAACTTCAAAGAACTGATCCTGGAAAACCACTACAACACCTACGCATCT-
 +---------+---------+---------+---------+---------+---------
 N  E  D  C  N  F  K  E  L  I  L  E  N  H  Y  N  T  Y  A  S -

-GCTAAATGGACCCACAACGGTGGTGAAATGTTCGTTGCTCTGAACCAGAAAGGTATCCCT-
 +---------+---------+---------+---------+---------+---------
 A  K  W  T  H  N  G  G  E  M  F  V  A  L  N  Q  K  G  I  P -

-GTTCTGGGTAAGAAAACCAAGAAAGAACAGAAAACCGCTCACTTCCTGCCGATGGCAATC-
 +---------+---------+---------+---------+---------+---------
 V  L  G  K  K  T  K  K  E  Q  K  T  A  H  F  L  P  M  A  I -

-ACTTAA-3'
 +-----
 T  *
```

Figure 22

ΔN23/K(147)E

```
5'-ATGTCCTACGACTACATGGAAGGTGGTGACATCCGCGTACGTCGTCTGTTCTGCCGTACC-
   +---------+---------+---------+---------+---------+---------
    M  S  Y  D  Y  M  E  G  G  D  I  R  V  R  R  L  F  C  R  T  -

-CAGTGGTACCTGCGTATCGACAAACGCGGCAAAGTCAAGGGCACCCAAGAGATGAAAAAC-
 +---------+---------+---------+---------+---------+---------
  Q  W  Y  L  R  I  D  K  R  G  K  V  K  G  T  Q  E  M  K  N  -

-AACTACAATATTATGGAAATCCGTACTGTTGCTGTTGGTATCGTTGCAATCAAAGGTGTT-
 +---------+---------+---------+---------+---------+---------
  N  Y  N  I  M  E  I  R  T  V  A  V  G  I  V  A  I  K  G  V  -

-GAATCTGAATTCTACCTGGCAATGAACAAAGAAGGTAAACTGTACGCAAAAAAAGAATGC-
 +---------+---------+---------+---------+---------+---------
  E  S  E  F  Y  L  A  M  N  K  E  G  K  L  Y  A  K  K  E  C  -

-AACGAAGACTGCAACTTCAAAGAACTGATCCTGGAAAACCACTACAACACCTACGCATCT-
 +---------+---------+---------+---------+---------+---------
  N  E  D  C  N  F  K  E  L  I  L  E  N  H  Y  N  T  Y  A  S  -

-GCTAAATGGACCCACAACGGTGGTGAAATGTTCGTTGCTCTGAACCAGAAAGGTATTCCT-
 +---------+---------+---------+---------+---------+---------
  A  K  W  T  H  N  G  G  E  M  F  V  A  L  N  Q  K  G  I  P  -

-GTTCGTGGTAAGGAAACCAAGAAAGAACAGAAAACCGCTCACTTCCTGCCGATGGCAATC-
 +---------+---------+---------+---------+---------+---------
  V  R  G  K  E  T  K  K  E  Q  K  T  A  H  F  L  P  M  A  I  -

-ACTTAA-3'
 +-----
  T  *
```

Figure 23

ΔN23/K(147)Q

```
5'-ATGTCCTACGACTACATGGAAGGTGGTGACATCCGCGTACGTCGTCTGTTCTGCCGTACC-
   +---------+---------+---------+---------+---------+---------
    M  S  Y  D  Y  M  E  G  G  D  I  R  V  R  R  L  F  C  R  T  -

  -CAGTGGTACCTGCGTATCGACAAACGCGGCAAAGTCAAGGGCACCCAAGAGATGAAAAAC-
   +---------+---------+---------+---------+---------+---------
    Q  W  Y  L  R  I  D  K  R  G  K  V  K  G  T  Q  E  M  K  N  -

  -AACTACAATATTATGGAAATCCGTACTGTTGCTGTTGGTATCGTTGCAATCAAAGGTGTT-
   +---------+---------+---------+---------+---------+---------
    N  Y  N  I  M  E  I  R  T  V  A  V  G  I  V  A  I  K  G  V  -

  -GAATCTGAATTCTACCTGGCAATGAACAAAGAAGGTAAACTGTACGCAAAAAAAGAATGC-
   +---------+---------+---------+---------+---------+---------
    E  S  E  F  Y  L  A  M  N  K  E  G  K  L  Y  A  K  K  E  C  -

  -AACGAAGACTGCAACTTCAAAGAACTGATCCTGGAAAACCACTACAACACCTACGCATCT-
   +---------+---------+---------+---------+---------+---------
    N  E  D  C  N  F  K  E  L  I  L  E  N  H  Y  N  T  Y  A  S  -

  -GCTAAATGGACCCACAACGGTGGTGAAATGTTCGTTGCTCTGAACCAGAAAGGTATCCCT-
   +---------+---------+---------+---------+---------+---------
    A  K  W  T  H  N  G  G  E  M  F  V  A  L  N  Q  K  G  I  P  -

  -GTTCGTGGTAAGCAAACCAAGAAAGAACAGAAAACCGCTCACTTCCTGCCGATGGCAATC-
   +---------+---------+---------+---------+---------+---------
    V  R  G  K  Q  T  K  K  E  Q  K  T  A  H  F  L  P  M  A  I  -

  -ACTTAA-3'
   +-----
    T  *
```

Figure 24

ΔN23/K(153)E

```
5'-ATGTCCTACGACTACATGGAAGGTGGTGACATCCGCGTACGTCGTCTGTTCTGCCGTACC-
   +---------+---------+---------+---------+---------+---------
    M   S   Y   D   Y   M   E   G   G   D   I   R   V   R   R   L   F   C   R   T  -

  -CAGTGGTACCTGCGTATCGACAAACGCGGCAAAGTCAAGGGCACCCAAGAGATGAAAAAC-
   +---------+---------+---------+---------+---------+---------
    Q   W   Y   L   R   I   D   K   R   G   K   V   K   G   T   Q   E   M   K   N  -

  -AACTACAATATTATGGAAATCCGTACTGTTGCTGTTGGTATCGTTGCAATCAAAGGTGTT-
   +---------+---------+---------+---------+---------+---------
    N   Y   N   I   M   E   I   R   T   V   A   V   G   I   V   A   I   K   G   V  -

  -GAATCTGAATTCTACCTGGCAATGAACAAAGAAGGTAAACTGTACGCAAAAAAAGAATGC-
   +---------+---------+---------+---------+---------+---------
    E   S   E   F   Y   L   A   M   N   K   E   G   K   L   Y   A   K   K   E   C  -

  -AACGAAGACTGCAACTTCAAAGAACTGATCCTGGAAAACCACTACAACACCTACGCATCT-
   +---------+---------+---------+---------+---------+---------
    N   E   D   C   N   F   K   E   L   I   L   E   N   H   Y   N   T   Y   A   S  -

  -GCTAAATGGACCCACAACGGTGGTGAAATGTTCGTTGCTCTGAACCAGAAAGGTATCCCT-
   +---------+---------+---------+---------+---------+---------
    A   K   W   T   H   N   G   G   E   M   F   V   A   L   N   Q   K   G   I   P  -

  -GTTCGTGGTAAGAAAACCAAGAAAGAACAGGAAACCGCTCACTTCCTGCCGATGGCAATC-
   +---------+---------+---------+---------+---------+---------
    V   R   G   K   K   T   K   K   E   Q   E   T   A   H   F   L   P   M   A   I  -

  -ACTTAA-3'
   +-----
    T   *
```

Figure 25

ΔN23/K(153)Q

```
5'-ATGTCCTACGACTACATGGAAGGTGGTGACATCCGCGTACGTCGTCTGTTCTGCCGTACC-
  +---------+---------+---------+---------+---------+---------
    M  S  Y  D  Y  M  E  G  G  D  I  R  V  R  R  L  F  C  R  T  -

- CAGTGGTACCTGCGTATCGACAAACGCGGCAAAGTCAAGGGCACCCAAGAGATGAAAAAC-
  +---------+---------+---------+---------+---------+---------
    Q  W  Y  L  R  I  D  K  R  G  K  V  K  G  T  Q  E  M  K  N  -

-AACTACAATATTATGGAAATCCGTACTGTTGCTGTTGGTATCGTTGCAATCAAAGGTGTT-
  +---------+---------+---------+---------+---------+---------
    N  Y  N  I  M  E  I  R  T  V  A  V  G  I  V  A  I  K  G  V  -

-GAATCTGAATTCTACCTGGCAATGAACAAAGAAGGTAAACTGTACGCAAAAAAAGAATGC-
  +---------+---------+---------+---------+---------+---------
    E  S  E  F  Y  L  A  M  N  K  E  G  K  L  Y  A  K  K  E  C  -

-AACGAAGACTGCAACTTCAAAGAACTGATCCTGGAAAACCACTACAACACCTACGCATCT-
  +---------+---------+---------+---------+---------+---------
    N  E  D  C  N  F  K  E  L  I  L  E  N  H  Y  N  T  Y  A  S  -

-GCTAAATGGACCCACAACGGTGGTGAAATGTTCGTTGCTCTGAACCAGAAAGGTATCCCT-
  +---------+---------+---------+---------+---------+---------
    A  K  W  T  H  N  G  G  E  M  F  V  A  L  N  Q  K  G  I  P  -

-GTTCGTGGTAAGAAAACCAAGAAAGAACAGCAAACCGCTCACTTCCTGCCGATGGCAATC-
  +---------+---------+---------+---------+---------+---------
    V  R  G  K  K  T  K  K  E  Q  Q  T  A  H  F  L  P  M  A  I  -

-ACTTAA-3'
  +-----
    T  *  -
```

Figure 26

ΔN23/Q(152)E/K(153)E

```
5'-ATGTCCTACGACTACATGGAAGGTGGTGACATCCGCGTACGTCGTCTGTTCTGCCGTACC-
  +---------+---------+---------+---------+---------+---------
   M  S  Y  D  Y  M  E  G  G  D  I  R  V  R  R  L  F  C  R  T  -

-CAGTGGTACCTGCGTATCGACAAACGCGGCAAAGTCAAGGGCACCCAAGAGATGAAAAAC-
 +---------+---------+---------+---------+---------+---------
  Q  W  Y  L  R  I  D  K  R  G  K  V  K  G  T  Q  E  M  K  N  -

-AACTACAATATTATGGAAATCCGTACTGTTGCTGTTGGTATCGTTGCAATCAAAGGTGTT-
 +---------+---------+---------+---------+---------+---------
  N  Y  N  I  M  E  I  R  T  V  A  V  G  I  V  A  I  K  G  V  -

-GAATCTGAATTCTACCTGGCAATGAACAAAGAAGGTAAACTGTACGCAAAAAAAGAATGC-
 +---------+---------+---------+---------+---------+---------
  E  S  E  F  Y  L  A  M  N  K  E  G  K  L  Y  A  K  K  E  C  -

-AACGAAGACTGCAACTTCAAAGAACTGATCCTGGAAAACCACTACAACACCTACGCATCT-
 +---------+---------+---------+---------+---------+---------
  N  E  D  C  N  F  K  E  L  I  L  E  N  H  Y  N  T  Y  A  S  -

-GCTAAATGGACCCACAACGGTGGTGAAATGTTCGTTGCTCTGAACCAGAAAGGTATCCCT-
 +---------+---------+---------+---------+---------+---------
  A  K  W  T  H  N  G  G  E  M  F  V  A  L  N  Q  K  G  I  P  -

-GTTCGTGGTAAGAAAACCAAGAAAGAAGAGGAAACCGCTCACTTCCTGCCGATGGCAATC-
 +---------+---------+---------+---------+---------+---------
  V  R  G  K  K  T  K  K  E  E  E  T  A  H  F  L  P  M  A  I  -

-ACTTAA-3'
 +-----
  T  *
```